# FINITE DIFFERENCES
## AND DIFFERENCE EQUATIONS IN THE REAL DOMAIN

BY

TOMLINSON FORT
PROFESSOR OF MATHEMATICS IN THE
UNIVERSITY OF GEORGIA

OXFORD
AT THE CLARENDON PRESS
1948

*Oxford University Press, Amen House, London E.C. 4*
EDINBURGH GLASGOW NEW YORK TORONTO MELBOURNE
WELLINGTON BOMBAY CALCUTTA MADRAS CAPE TOWN
*Geoffrey Cumberlege, Publisher to the University*

PRINTED IN GREAT BRITAIN

# PREFACE

THIS volume includes material for a course in Finite Differences as well as a treatment of a number of special topics. A selection of material was necessary in order to bring the work within reasonable compass and also to follow somewhat the special interests of the author. The very large theory of difference equations in the field of analytic functions of a complex variable has been completely omitted. This has carried with it the omission of a variety of related subjects, particularly in Infinite Series.

Material for the book has been taken from any available source. There is much material that hitherto has not been published. A great deal of this represents research of the author. Some was taken from interpolations in lectures on linear differential equations delivered years ago by Bôcher. His influence is strong in Chapters IX, X, and the first part of XII.

The author believes that the treatment of the linear recurrent relation, that is, of the linear difference equation with independent variable limited to integral values, will be found particularly interesting. In fact this constitutes the major part of the book and is the portion in which most of the original work is located. Chapter XVII is really a supplement to Chapter IX and if the book is used as a basis for a course of lectures the material in Chapter XVII might well follow that in Chapter IX.

When the preparation of this book was already well under way Professor J. A. Shohat, of the University of Pennsylvania, was associated with the present writer, and it was planned to bring out a book under joint authorship. However, increase in duties due to the War and rapidly failing health allowed Professor Shohat to make but one substantial contribution, namely, a first draft of Chapter VI on 'Interpolation and Mechanical Quadratures'. Although this chapter has been completely rewritten it follows closely the original draft of Professor Shohat.

Not a great many references are given. A few papers, mostly well-known ones, are cited, and the reader is referred at the end of some chapters to certain pages in standard reference books

for the reading of related material. At the end of the volume some books are listed which can be used for further study. No attempt is made to give a complete bibliography. For a list of books and papers up to the year 1930 the reader is referred to the 'Literatur Verzeichnis' given by N. E. Nörlund in his *Differenzenrechnung*, also to the 'Bibliography' of C. R. Adams in the *Bulletin of the American Mathematical Society*, volume 37, page 383.

T. F.

# CONTENTS

I

# DIRECT DIFFERENCE OPERATORS

## 1. The operator $\Delta$

In the present chapter we shall write $u(x)$, $v(x)$, etc., and make no assumptions about these functions beyond the fact that they are defined for all values of the independent variable considered.

Let $h$ be a positive constant which we shall call the difference interval. Then, by definition,

$$\begin{aligned} \Delta u(x) &= u(x+h)-u(x), \\ \Delta^2 u(x) &= \Delta\{\Delta u(x)\}, \\ &\cdots\cdots\cdots \\ \Delta^n u(x) &= \Delta\{\Delta^{n-1}u(x)\}. \end{aligned} \tag{1}$$

These are called respectively the first, second, and $n$th differences of $u(x)$.

Theorem I.

$$\begin{aligned} \Delta^n u(x) = {} & u(x+nh)-nu\{x+(n-1)h\}+ \\ & +\frac{n(n-1)}{2!}u\{x+(n-2)h\}-\dots+ \\ & +(-1)^k\frac{n(n-1)\dots(n-k+1)}{k!}u\{x+(n-k)h\}+\dots+(-1)^n u(x). \end{aligned} \tag{2}$$

This theorem is readily proved by mathematical induction. Write formula (2), replacing $x$ by $x+h$,

$$\begin{aligned} & \Delta^n u(x+h) \\ & = u\{x+(n+1)h\}-nu(x+nh)+\frac{n(n-1)}{2!}u\{x+(n-1)h\}- \\ & \qquad -\dots+(-1)^{k+1}\frac{n(n-1)\dots(n-k)}{(k+1)!}u\{x+(n-k)h\}+\dots+ \\ & \qquad +(-1)^n u(x+h). \end{aligned}$$

Then

$$\begin{aligned}\Delta^{n+1}u(x) &= \Delta^n u(x+h)-\Delta^n u(x)\\ &= u\{x+(n+1)h\}-(n+1)u(x+nh)+\frac{(n+1)n}{2!}u\{x+(n-1)h\}-\\ &\quad -\ldots+(-1)^{k+1}\frac{(n+1)n(n-1)\ldots(n-k+1)}{(k+1)!}u\{x+(n-k)h\}+\\ &\quad +\ldots+(-1)^{n+1}u(x).\end{aligned}$$

This is of the same form as (2) with $n$ replaced by $(n+1)$, and induction is immediate, since (2) holds for $n = 1$.

The inverse problem to Theorem I is covered by the following.

THEOREM II.

$$u(x+nh) = u(x)+n\Delta u(x)+\frac{n(n-1)}{2!}\Delta^2 u(x)+\ldots+\Delta^n u(x). \tag{3}$$

Proof of this formula by mathematical induction can be carried out as in the previous case.

Formula (2) can be written

$$\Delta^n u(x) = \sum_{i=0}^{n}(-1)^i {}_nC_{n-i}\, u\{x+(n-i)h\}, \tag{4}$$

where ${}_nC_{n-i}$ is a binomial coefficient.

If now we note that

$$0 = (1-1)^n = \sum_{i=0}^{n}(-1)^i {}_nC_{n-i}$$

we have immediately the following theorem:

THEOREM III.

$$\Delta^n u(x) = \sum_{i=0}^{n-1}(-1)^i {}_nC_{n-i}[u\{x+(n-i)h\}-u(x)]. \tag{5}$$

## 2. Analogues of simple formulae from differential calculus

The following laws governing the operator $\Delta$ follow directly from its definition.

If $c(x+h) = c(x)$, and if no ambiguity is involved, we shall refer to $c(x)$ as a constant and simply write $c$.

$$\Delta c = 0. \tag{6}$$

$$\Delta cu(x) = c\Delta u(x). \tag{7}$$

$$\Delta\{u(x)+v(x)\} = \Delta u(x)+\Delta v(x). \tag{8}$$

$$\Delta^m\{\Delta^n u(x)\} = \Delta^{m+n}u(x). \tag{9}$$

$$\Delta u(x)v(x) = u(x)\Delta v(x)+v(x+h)\Delta u(x). \tag{10}$$

$$\Delta\frac{u(x)}{v(x)} = \frac{v(x)\Delta u(x)-u(x)\Delta v(x)}{v(x+h)v(x)}; \qquad v(x+h)v(x) \neq 0. \tag{11}$$

$$\Delta[x(x-h)...\{x-(n-1)h\}] = nh[x(x-h)...\{x-(n-2)h\}]. \tag{12}$$

We call $x(x-h)...\{x-(n-1)h\}$ a 'factorial' of degree $n$ and denote it by $x^{(n)}$.

We can write (12) as

$$\Delta x^{(n)} = nhx^{(n-1)}. \tag{12$'$}$$

Similarly, if we let $x^{(-n)} = \dfrac{1}{x(x+h)...\{x+(n-1)h\}}$, we have

$$\Delta x^{(-n)} = -nhx^{\{-(n+1)\}}, \tag{13}$$

$$\Delta[u(x)u(x-h)...u\{x-(n-1)h\}]$$
$$= [u(x+h)-u\{x-(n-1)h\}]u(x)u(x-h)...u\{x-(n-2)h\}, \tag{14}$$

$$\Delta\frac{1}{u(x)u(x+h)...u\{x+(n-1)h\}} = \frac{u(x)-u(x+nh)}{u(x)u(x+h)...u(x+nh)}. \tag{15}$$

In particular if $u(x) = ax+b$, (14) and (15) become

$$\Delta[u(x)u(x-h)...u\{x-(n-1)h\}]$$
$$= anhu(x)u(x-h)...u\{x-(n-2)h\}, \tag{14$'$}$$

$$\Delta\frac{1}{u(x)u(x+h)...u\{x+(n-1)h\}} = \frac{-anh}{u(x)u(x+h)...u(x+nh)}. \tag{15$'$}$$

$$\Delta\log_a x = \log_a\left(1+\frac{h}{x}\right). \tag{16}$$

$$\Delta\log_a u(x) = \log_a\left(1+\frac{\Delta u(x)}{u(x)}\right). \tag{16$'$}$$

$$\Delta C^x = (C^h-1)C^x. \tag{17}$$

$$\Delta^n C^x = (C^h-1)^n C^x. \tag{17$'$}$$

$$\Delta\sin(ax+b) = 2\sin\tfrac{1}{2}ah\cos(ax+b+\tfrac{1}{2}ah). \tag{18}$$

$$\Delta\cos(ax+b) = -2\sin\tfrac{1}{2}ah\sin(ax+b+\tfrac{1}{2}ah). \tag{19}$$

$$\Delta \tan x = \sec^2 x \frac{\tan h}{1-\tan h \tan x}. \tag{20}$$

$$\Delta \tan^{-1} u(x) = \tan^{-1} \frac{\Delta u(x)}{1+u(x+h)u(x)}. \tag{21}$$

$$\Delta \sec x = -\frac{\Delta \cos x}{\cos(x+h)\cos x}. \tag{22}$$

$$\Delta \sinh(ax+b) = 2\sinh \tfrac{1}{2}ah \cosh(ax+b+\tfrac{1}{2}ah). \tag{23}$$

$$\Delta \cosh(ax+b) = 2\sinh \tfrac{1}{2}ah \sinh(ax+b+\tfrac{1}{2}ah). \tag{24}$$

The above list of formulae is clearly in no way exhaustive. In fact, ordinarily we expect to be able to write a useful difference formula for any function $u(x)$ where we can express $u(x+h)$ in a simple manner in terms of $u(x)$ and constants.

## 3. Table of differences

The finding of successive differences for a given function is frequently facilitated by the construction of a table as follows:

$$\begin{array}{llllll}
x & u(x) & & & & \\
 & & \Delta u(x) & & & \\
x+h & u(x+h) & & \Delta^2 u(x) & & \\
 & & \Delta u(x+h) & & \Delta^3 u(x) & \\
x+2h & u(x+2h) & & \Delta^2 u(x+h) & & \\
 & & \Delta u(x+2h) & & \Delta^3 u(x+h) & \\
x+3h & u(x+3h) & & \Delta^2 u(x+2h) & & \\
 & & \Delta u(x+3h) & & & \\
x+4h & u(x+4h) & & & & \\
\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot & .\,\Delta^n u(x) \\
x+(n-1)h & u\{x+(n-1)h\} & & & & \\
 & & \Delta u\{x+(n-1)h\} & & & \\
x+nh & u(x+nh) & & & &
\end{array}$$

Each entry in the table beginning with the third column is the difference of the two entries in the column immediately preceding it. Note that in order to evaluate $\Delta^n u(x)$ we need the functional values $u(x)$, $u(x+h)$,..., $u(x+nh)$.

## 4. A further difference formula

Observe next formula (8). This can be described by saying that *the operator* $\Delta$ *is distributive with respect to addition.* Formula (7) states that $\Delta$ *is commutative with respect to a constant.*

We describe formula (9) by saying that $\Delta$ *obeys the index law.* Now suppose that we are given two polynomials in a variable $z$,

$$f_1(z) \equiv a_0 z^n + a_1 z^{n-1} + \dots + a_n,$$
$$f_2(z) \equiv b_0 z^m + b_1 z^{m-1} + \dots + b_m.$$

Form the product of these polynomials,

$$F(z) \equiv f_1(z) f_2(z) = c_0 z^{m+n} + c_1 z^{m+n-1} + \dots + c_{m+n}.$$

Now build up the operators,

$$f_1(\Delta) = a_0 \Delta^n + a_1 \Delta^{n-1} + \dots + a_n,$$
$$f_2(\Delta) = b_0 \Delta^m + b_1 \Delta^{m-1} + \dots + b_m,$$
$$F(\Delta) = c_0 \Delta^{m+n} + c_1 \Delta^{m+n-1} + \dots + c_{m+n}.$$

The meaning of these operators is explained by the following identical relation employing $f_1(\Delta)$,

$$f_1(\Delta)u(x) \equiv a_0 \Delta^n u(x) + a_1 \Delta^{n-1} u(x) + \dots + a_n u(x).$$

Notice that the index law, the distributive law with reference to addition, and the commutative law with reference to constants are the sole laws used in forming the polynomial $F$ from the polynomials $f_1$ and $f_2$ by multiplication. It follows that

$$f_1(\Delta)[f_2(\Delta)u(x)] \equiv f_2(\Delta)[f_1(\Delta)u(x)] \equiv F(\Delta)u(x). \tag{25}$$

It is instructive to carry through the actual operations and thus verify (25) for two polynomials of low degree.

We can readily prove formula (3) by means of (25). We see immediately that

$$u(x+h) = (1+\Delta)u(x).$$

Whereupon

$$u(x+nh) = (1+\Delta)^n u(x). \tag{26}$$

## 5. The operator $E$

It is frequently convenient to introduce an operator $E$ defined by the following relations†

$$Eu(x) = (1+\Delta)u(x) = u(x+h),$$
$$E^n u(x) = u(x+nh),$$

so that

$$\Delta u(x) = (E-1)u(x).$$

† In the definition of $E^n$ it is not necessary that $n$ be a positive integer, although this will be done in the sequel unless the contrary is explicitly stated.

We have already given the important formula (26). It is interesting to note that this can be written using the symbol $E$,

$$E^n u(x) = (1+\Delta)^n u(x). \tag{27}$$

Laws of operation with $E$ are as follows:

$$Ecu(x) = cEu(x),$$
$$E\{u(x)+v(x)\} = Eu(x)+Ev(x),$$
$$E^m[E^n u(x)] = E^{m+n}u(x) = E^n[E^m u(x)].$$

Moreover if $f_1(z)$ and $f_2(z)$ are polynomials in $z$ as defined in § 4, then

$$f_1(E)[f_2(E)u(x)] = f_2(E)[f_1(E)u(x)] = F(E)u(x)$$
$$= c_0\, E^{m+n}u(x)+c_1\, E^{m+n-1}u(x)+\ldots+c_{m+n}\, u(x). \tag{28}$$

This follows exactly as in the case of the operator $\Delta$.

Formula (2) can now be readily proved by means of the binomial theorem. We simply write

$$\Delta^n u(x) = (E-1)^n u(x). \tag{29}$$

## 6. The $n$th difference of a product

An interesting formula is the following:

$$\Delta^n\{u(x)v(x)\} = \{\Delta^n u(x)\}v(x+nh)+n\{\Delta^{n-1}u(x)\}\Delta v\{x+(n-1)h\}+$$
$$+{}_nC_2\{\Delta^{n-2}u(x)\}\Delta^2 v\{x+(n-2)h\}+\ldots+u(x)\Delta^n v(x). \tag{30}$$

Proof of this by mathematical induction can be carried out. However, proof by means of the symbols $E$ and $\Delta$ is brief and proceeds as follows:

Let $E''$ and $\Delta''$ be applied only to $u(x)$ and $E'$ and $\Delta'$ to $v(x)$. Otherwise the accent is without significance. Then,

$$\Delta\{u(x)v(x)\} = u(x+h)v(x+h)-u(x)v(x) = (E''E'-1)u(x)v(x),$$
$$\Delta^n\{u(x)v(x)\} = (E''E'-1)^n u(x)v(x) = \{(1+\Delta'')E'-1\}^n u(x)v(x)$$
$$= (E'\Delta''+\Delta')^n u(x)v(x).$$

Expand by the binomial theorem and remember that $\Delta'$ and $E'$ operate only on $v(x)$ and $\Delta''$ on $u(x)$, then drop the accents from $\Delta$ as being no longer relevant, and we have the required formula.

## 7. Differencing rational functions

Any polynomial can be written as a sum of factorials, thus,

$$\begin{aligned} x^2 &= x(x-h)+hx, \\ x^3 &= x(x-h)(x-2h)+3hx^2-2h^2x \\ &= x(x-h)(x-2h)+3h[x(x-h)+hx]-2h^2x \\ &= x^{(3)}+3hx^{(2)}+h^2x^{(1)}. \end{aligned}$$

In general, if

$$f(x) = a_0+a_1x+a_2x^2+...+a_nx^n,$$

it is possible to write

$$f(x) = A_0+A_1x^{(1)}+A_2x^{(2)}+...+A_nx^{(n)},$$

where the $A$'s are independent of $x$. Let $x = 0$ and we find $A_0$. Take the difference of both sides and again let $x = 0$.

We get $A_1 = \frac{1}{h}\Delta f(0)$, and in general $A_i = \frac{1}{h^i}\frac{\Delta^i f(0)}{i!}$, $i = 0$, $1,..., n$. Here $\Delta^0 f(0) = f(0)$ and $0! = 1$.

Hence,

$$f(x) = f(0)+\frac{1}{h}\Delta f(0)x^{(1)}+\frac{1}{h^2}\Delta^2 f(0)\frac{x^{(2)}}{2!}+...+\frac{1}{h^n}\Delta^n f(0)\frac{x^{(n)}}{n!}. \tag{31}$$

This is an identity. Formula (31) is widely used as an interpolation formula under the name *Newton's Interpolation Formula.* If throughout its derivation we replace $x$ by $x-a = k$, we get the more general formula

$$f(x) = f(a)+\frac{1}{h}\Delta f(a)(x-a)^{(1)}+\frac{1}{h^2}\Delta^2 f(a)\frac{(x-a)^{(2)}}{2!}+...+ \\ +\frac{1}{h^n}\Delta^n f(a)\frac{(x-a)^{(n)}}{n!}. \tag{32}$$

This can be written

$$f(a+k) = f(a)+\frac{1}{h}\Delta f(a)k^{(1)}+\frac{1}{h^2}\Delta^2 f(a)\frac{k^{(2)}}{2!}+...+\frac{1}{h^n}\Delta^n f(a)\frac{k^{(n)}}{n!}.$$

We can now find successive differences of $f(x)$ using formulae (31) and (32). For example,

$$\Delta f(x) = \Delta f(0)+\frac{1}{h}\Delta^2 f(0)x^{(1)}+...+\frac{1}{h^{n-1}}\Delta^n f(0)x^{(n-1)}.$$

In case $f(x)$ is a rational fraction it may happen that formula (13) for $\Delta x^{(-n)}$ is helpful.

Methods will be illustrated by an example: Suppose $h = 1$ and

$$f(x) = \frac{x^2+8x+16}{x(x+1)(x+2)}.$$

By the method of 'undetermined coefficients'

$$f(x) = x^{(-1)}+5x^{(-2)}+4x^{(-3)}.$$

From this with the use of (13)

$$\Delta f(x) = -x^{(-2)}-10x^{(-3)}-12x^{(-4)}.$$

By inspection

$$x^2+8x+16 = 16+7x+x(x+1),$$

whence $$f(x) = 16x^{(-3)}+7(x+1)^{(-2)}+(x+2)^{(-1)},$$

which can be differenced term by term. We get

$$\Delta f(x) = -48x^{(-4)}-14(x+1)^{(-3)}-(x+2)^{(-2)}.$$

Similarly, separating $f(x)$ into partial fractions, we can write

$$\frac{x^2+8x+16}{x(x+1)(x+2)} = 8x^{-1}-9(x+1)^{-1}+2(x+2)^{-1}.$$

Whereupon

$$\Delta f(x) = -8x^{(-2)}+9(x+1)^{(-2)}-2(x+2)^{(-2)}.$$

The difference of this function also can be readily found if we write it as a product and use the product-rule formula (10). Thus

$$f(x) = (x^2+8x+16)x^{(-3)}.$$

We also can use the fraction-rule formula (11).

There is in addition always the possibility of working out directly $f(x+h)-f(x)$.

## 8. Backward differences

It is occasionally convenient to introduce 'backward differences' defined by the formulae

$$\underline{\Delta} u(x) = -u(x)+u(x-h),$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\underline{\Delta}^n u(x) = \underline{\Delta}\{\underline{\Delta}^{n-1}u(x)\}.$$

It is easily shown that

$$\underline{\Delta}^n u(x) = (-1)^n \Delta^n u(x-nh).$$

## 9. Difference quotients

In much work in finite differences it is convenient to deal with the difference quotients:

$$\underset{h}{\Delta} u(x) = \frac{u(x+h)-u(x)}{h},$$

$$\cdots\cdots\cdots$$

$$\underset{h}{\Delta}^n u(x) = \frac{\Delta^n u(x)}{h^n}.$$

If $h = 1$ $$\underset{h}{\Delta} u(x) = \Delta u(x).$$

Formulae governing operation with $\underset{h}{\Delta}$ are readily inferred from § 2.

## 10. Central differences

Let

$$\delta u(x) \equiv u(x+\tfrac{1}{2}h)-u(x-\tfrac{1}{2}h)$$
$$\equiv (E^{\frac{1}{2}}-E^{-\frac{1}{2}})u(x) \equiv \Delta(x-\tfrac{1}{2}h).$$

Then $\delta u(x)$ is known as the *central difference of* $u(x)$ with difference interval $h$.

Inasmuch as the central difference has been expressed by means of the symbol $E$ and by means of the symbol $\Delta$, rules of operation can be deduced from the rules for operating with $E$ and $\Delta$.

## 11. The operator $\nabla$, the mean

This operator is defined by

$$\nabla u(x) \equiv \tfrac{1}{2}[u(x+h)+u(x)] \equiv \tfrac{1}{2}[E+1]u(x),$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\nabla^n u(x) = \nabla[\nabla^{n-1}u(x)].$$

It is immediate that

$$\nabla^n u(x) = \frac{1}{2^n}[E+1]^n u(x).$$

## 12. Divided differences

By definition

$$\eta u(x) = \frac{u(x+h)-u(x)}{h},$$

$$\eta^2 u(x) = \frac{\eta u(x+h)-\eta u(x)}{2h},$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$\eta^n u(x) = \frac{\eta^{n-1}u(x+h)-\eta^{n-1}u(x)}{nh}.$$

These are called the first, second, and $n$th divided differences of $u(x)$. They have proved particularly useful in interpolation problems. *There is no generally accepted notation for divided difference.*

The notion of divided differences is readily generalized to unequally spaced points. Let $x_1,\ldots, x_n$ be $n$ points. Then we write

$$[x_2 x_1] = \frac{u(x_2)-u(x_1)}{x_2-x_1},$$

$$[x_3 x_2 x_1] = \frac{[x_3 x_2]-[x_2 x_1]}{x_3-x_1},$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$[x_n x_{n-1} \ldots x_1] = \frac{[x_n x_{n-1} \ldots x_2]-[x_{n-1} x_{n-2} \ldots x_1]}{x_n-x_1}.$$

Then $[x_j x_{j-1} \ldots x_1]$ is called the $j$th divided difference of $u(x)$ with reference to $x_1,\ldots, x_n$. The notation has the disadvantage of not explicitly writing $u(x)$.

THEOREM IV.

$$[x_n x_{n-1} \ldots x_1] = \frac{u(x_1)}{(x_1-x_2)(x_1-x_3)\ldots(x_1-x_n)}+$$
$$+\frac{u(x_2)}{(x_2-x_1)(x_2-x_3)\ldots(x_2-x_n)}+\ldots+\frac{u(x_n)}{(x_n-x_1)(x_n-x_2)\ldots(x_n-x_{n-1})}. \tag{33}$$

This is easily proved by mathematical induction.

From formula (33) we immediately infer the following theorem.

THEOREM V. $[x_n\, x_{n-1} \dots x_1]$ *is symmetrical in the arguments*

$$x_1, \dots, x_n.$$

Also from (33) we conclude the following theorem.

THEOREM VI.

$$[x_n\, x_{n-1} \dots x_1] = \frac{\begin{vmatrix} 1 & x_1 & . & . & . & x_1^{n-1} & u(x_1) \\ 1 & x_2 & . & . & . & x_2^{n-1} & u(x_2) \\ . & . & . & . & . & . & . \\ 1 & x_n & . & . & . & x_n^{n-1} & u(x_n) \end{vmatrix}}{\begin{vmatrix} 1 & x_1 & . & . & . & x_1^{n-1} & x_1^n \\ 1 & x_2 & . & . & . & x_2^{n-1} & x_2^n \\ . & . & . & . & . & . & . \\ 1 & x_n & . & . & . & x_n^{n-1} & x_n^n \end{vmatrix}}.$$

Divided differences have been extensively studied. However, only the briefest facts are given here. We state a few under one theorem heading. Some additional theorems are given as exercises.

THEOREM VII. *A change of origin does not affect the divided difference.*

*The operation of taking the divided difference is associative with reference to addition: it is commutative with reference to multiplication by a constant.*

*If $u(x)$ is a polynomial of the n-th degree $[x_1\, x]$ is of the $(n-1)$th degree.*

Newton's formula permits of immediate extension to divided differences. This is particularly valuable if the points considered are unequally spaced.

Let $u(x)$ be a polynomial of degree $n$ and let there be given a set of numbers $x_1, \dots, x_n$. Clearly we can write

$$u(x) = A_0 + A_1(x-x_1) + A_2(x-x_1)(x-x_2) + \\ + \dots + A_n(x-x_1)(x-x_2)\dots(x-x_n).$$

We readily determine the coefficients:

$$A_0 = u(x_1), \quad A_1 = [x_2\, x_1], \quad A_2 = [x_3\, x_2\, x_1], \quad \dots, \\ A_n = [x_n\, x_{n-1} \dots x_1].$$

## EXERCISES

1. Write $F(x)$ as a sum of the factorials $x^{(3)}$, $x^{(2)}$, $x^{(1)}$, $x^{(0)}$, and find $\Delta F(x)$, where

   (a) $F(x) = 3x^3+4x^2+7x-2$,

   (b) $F(x) = 7x^3+3x^2+8x-2$.

2. Write $F(x)$ as a sum of the factorials
$$(x-2)^{(3)},\quad (x-2)^{(2)},\quad (x-2)^{(1)},\quad (x-2)^{(0)},$$
and find $\Delta F(x)$, where $F(x)$ is as in Exercise 1.

3. Find $\Delta F(x)$ if $F(x)$ equals

   (a) $4x^{(-3)}+7x^{(-2)}+4$,

   (b) $x^{(-5)}x^{(7)}$,

   (c) $\sin 3x$,

   (d) $\sec(5x+6)$,

   (e) $10^{8x-3}$.

4. Find $\Delta^5(\sin 2x)10^{3x}$.

5. Using a table of trigonometric functions find

   (a) $\Delta^4\sin 30^\circ$, $h = 5^\circ$.

   (b) $\delta^2\sin 30^\circ$, $h = 10^\circ$.

   (c) $\nabla^3\sin 30^\circ$, $h = 5^\circ$.

   (d) $\eta^3\sin 30^\circ$, $h = 5^\circ$.

6. Find $\Delta^2\dfrac{3x+2}{(x+5)(x+6)(x+7)}$, $h = 1$.

7. Find $\underset{-}{\Delta}^3\dfrac{3x+2}{(x+5)(x+6)(x+7)}$, $h = 1$.

8. Derive formulae for $\Delta^n\sin x$ and $\Delta^n\cos x$.

9. Derive a formula for $\Delta^n[e^x u(x)]$.

10. Prove that $F(E)a^x\phi(x) = a^x F(a^h E)\phi(x)$, where $F(x)$ and $\phi(x)$ are polynomials.

11. Prove
$$\frac{1}{x^2+nx} = x^{(-2)}+(1-n)x^{(-3)}+(1-n)(2-n)x^{(-4)}+\ldots+$$
$$+(1-n)(2-n)\ldots(-1)x^{(-(n+1))}.$$

12. Given $u(x) = \dfrac{1}{x^2+5x}$ find $\Delta u(x)$ utilizing the formula of Exercise 11.

13. If $\phi(x)$ is a polynomial, prove
$$\phi(E)u(x) = a^x\phi(a^h E)\{a^{-x}u(x)\}.$$

14. Prove that $\Delta^n\nabla^n u(x)$, with the difference interval $h$, equals $\dfrac{1}{2^n}\Delta^n u(x)$, with the difference interval $2h$.

15. If $x_0$, $x_1$, $x_2$,..., $x_r$ are equally spaced, prove that

$$[x_r\, x_{r-1} \dots x_0] = \frac{1}{r!}\,\frac{1}{h^r}\Delta^r u(x_0).$$

16. Prove the formula:

$$[x_r\, x_{r-1} \dots x_0] = \int\limits_0^1 \int\limits_0^{t_1} \dots \int\limits_0^{t_{r-1}} u^{(r)}\left\{x_0\, t_r + \sum_{\nu=1}^{r} x_\nu(1-t_\nu)\right\} dt_r \dots dt_2\, dt_1.$$

17. If $u(x) = 1/x$ find $[x_r\, x_{r-1} \dots x_1]$.

II

# ELEMENTARY THEORY OF SUMMATION

## 1. Indefinite summation

IF $\Delta\phi(x) = \phi(x+h)-\phi(x) = f(x)$ *then, by definition,*

$$\phi(x) = \sum f(x).$$

$\sum f(x)$ is called the indefinite sum of $f(x)$ or where no ambiguity is involved simply the sum of $f(x)$. We shall speak of summing $f(x)$ with obvious import.

THEOREM I.

$$\Delta\phi(x) = f(x),$$
$$\Delta\Phi(x) = f(x),$$

*then* $\Phi(x)-\phi(x)$ *is a function with period* $h$.

This follows immediately from the fact that

$$\Delta\{\Phi(x)-\phi(x)\} = 0.$$

It is frequently convenient to refer to functions of period $h$ as constants. As a matter of fact, in much of our work the independent variable is limited to a set of discrete values with difference interval $h$. In this circumstance this function is truly a constant.

Other immediate theorems on summation are expressed by the formulae:

$$\sum\{f_1(x)+f_2(x)\} = \sum f_1(x)+\sum f_2(x), \tag{1}$$

$$\sum cf(x) = c\sum f(x), \quad c \text{ a constant}, \tag{2}$$

$$\sum u(x)\Delta v(x) = u(x)v(x)-\sum v(x+h)\Delta u(x), \tag{3}$$

$$\sum u(x+h)\Delta v(x) = u(x)v(x)-\sum v(x)\Delta u(x). \tag{4}$$

Formulae (3) and (4) are called *summation by parts formulae.*

In fact a summation formula corresponds to every difference formula. A few additional formulae will be given:

$$\sum x^{(n)} = \frac{1}{(n+1)h}x^{(n+1)}+c, \quad n > 0. \tag{5}$$

$$\sum x^{(-n)} = -\frac{1}{(n-1)h}x^{(-(n-1))}+c, \quad n > 1. \tag{6}$$

$$\sum m^x = \frac{m^x}{m^h-1}+c, \quad |m| \neq 1. \tag{7}$$

$$\sum \sin x = -\frac{\cos(x-\frac{1}{2}h)}{2\sin\frac{1}{2}h}+c. \tag{8}$$

To derive (8), note that

$$\Delta \cos x = -2\sin\tfrac{1}{2}h\sin(x+\tfrac{1}{2}h),$$

$$\Delta \cos(x-\tfrac{1}{2}h) = -2\sin\tfrac{1}{2}h\sin x$$

and sum both sides.

Similarly,

$$\sum \cos x = \frac{\sin(x-\frac{1}{2}h)}{2\sin\frac{1}{2}h}+c, \tag{9}$$

$$\sum \sinh x = \frac{\cosh(x-\frac{1}{2}h)}{2\sinh\frac{1}{2}h}+c. \tag{10}$$

As an application of these formulae consider the following:

$$\sum x^2 = \sum [hx+x^{(2)}] = \tfrac{1}{2}x^{(2)}+\frac{1}{3h}x^{(3)}+c. \tag{11}$$

In general, if $f(x)$ is a polynomial write it as a sum of factorials and sum.

If $f(x)$ is a rational fraction, devices such as separation into partial fractions and summation by parts may be helpful, although no certain rule of procedure can be given for obtaining a closed form for the sum in terms of the so-called elementary functions. In fact many sums cannot be so expressed. This is true of a sum so simple in appearance as $\sum 1/x$. We note in passing the following two additional formulae: if $u(x) = ax+b$, then

$$\sum u(x)u(x-h)...u\{x-(n-1)h\}$$
$$= \frac{1}{(n+1)ah}u(x)u(x-h)...u(x-nh)+c, \tag{12}$$

$$\sum \frac{1}{u(x)u(x+h)...u\{x+(n-1)h\}}$$
$$= \frac{-1}{(n-1)ahu(x)u(x+h)...u\{x+(n-2)h\}}+c. \tag{13}$$

By means of these formulae we obtain, for example ($h = 1$),

$$\sum (2x+5)(2x+3)(2x+1) = \frac{(2x+5)(2x+3)(2x+1)(2x-1)}{2.4}+c,$$

$$\sum \frac{1}{(3x-2)(3x+1)(3x+4)} = -\frac{1}{3.2(3x-2)(3x+1)}+c.$$

## 2. Summation by parts

The problem of summing a product of two functions is a close parallel to the corresponding problem in the integral calculus. Elementary examples where summation by parts is useful are as follows:

$$\sum xa^x = \frac{1}{a^h-1}xa^x-\frac{1}{a^h-1}\sum a^{x+h}h = \frac{1}{a^h-1}xa^x-\frac{ha^{x+h}}{(a^h-1)^2}+c, \tag{14}$$

$$\sum x\sin x = \frac{-x\cos(x-\frac{1}{2}h)}{2\sin\frac{1}{2}h}+\frac{h}{2\sin\frac{1}{2}h}\sum\cos(x+\tfrac{1}{2}h)$$

$$= \frac{-x\cos(x-\frac{1}{2}h)}{2\sin\frac{1}{2}h}+\frac{h}{(2\sin\frac{1}{2}h)^2}\sin x+c. \tag{15}$$

## 3. Definite summation

The topic of the previous section will be greatly elucidated by the discussion of definite summation which is introduced at this point.

By definition†

$$u(a)+u(a+h)+...+u(x-h) \equiv \sum_{\nu=a}^{x-h} u(\nu)$$

$$= \sum_{i=0}^{y} u(a+ih), \quad i = 0,1,...,y, \quad y = \frac{x-h-a}{h}. \tag{16}$$

Let

$$\phi(x) = \sum_{\nu=a}^{x-h} u(\nu), \quad x > a,$$

$$\phi(x) = 0, \quad x = a.$$

Then it is immediate that

$$\Delta\phi(x) = u(x), \quad x \geqslant a.$$

† The variable of summation will be denoted by the same letter as a limit variable if it seems to make for simplicity, thus

$$\sum_{x=a}^{x-h} u(x) \equiv \sum_{\nu=a}^{x-h} u(\nu).$$

Consequently, $\phi(x)$ is a particular determination of

$$\sum u(x); \quad x = a+\nu h, \quad \nu = 0, 1, 2,\dots.$$

If $u(x)$ is defined for all values of $x$ we can still use essentially the same scheme. Choose $\alpha$, so that

$$x = \alpha+ih, \quad i = 0, \pm 1, \pm 2,\dots.$$

Let

$$\phi(x) = \sum_{\nu=\alpha}^{x-h} u(\nu) \quad \text{when } x > \alpha \text{ as before,}$$

$$\phi(\alpha) = 0,$$

and

$$\phi(x) = -\sum_{\nu=x}^{\alpha-h} u(\nu)$$

$$= -u(x)-u(x+h)-\dots-u(\alpha-h) \quad \text{when } x < \alpha.$$

Again $$\Delta\phi(x) = u(x).$$

By subtraction we get the general formula

$$\sum_{\nu=b}^{c-h} u(\nu) = \phi(c)-\phi(b), \quad b = c+ih,$$

$i$ being an integer.

This formula holds for any determination of $\sum u(x)$, inasmuch as such a determination can at most differ from $\phi(x)$ by a constant (a function with period $h$). We write

$$\sum_{x=b}^{c-h} u(x) = \sum u(x)\Big]_c - \sum u(x)\Big]_b, \tag{17}$$

where $\sum u(x)$ is some particular determination.

Formula (17) can be used in the evaluation of definite sums. For example, with $h = 1$, from (11)

$$\sum_{x=1}^{n} x^2 = \left[\tfrac{1}{2}x(x-1)+\tfrac{1}{3}x(x-1)(x-2)\right]_1^{n+1}$$
$$= \tfrac{1}{2}(n+1)n+\tfrac{1}{3}(n+1)n(n-1).$$

A definite sum of any polynomial can be found in like manner by first expressing it as a sum of factorials.

Similarly, from (12)

$$\sum_{x=1}^{n} (2x+5)(2x+3)(2x+1)$$
$$= \frac{(2n+7)(2n+5)(2n+3)(2n+1)-7.5.3.1}{2.4}.$$

The finding of closed formulae like the above for finite sums is often not possible. The problem is analogous to the familiar one in integration.

## 4. The function $\underset{h}{L}(x, a)$

The problem of integration in ordinary calculus is greatly facilitated by the introduction of $\log x$ which is best defined by the formula

$$\int_1^x \frac{1}{x}\,dx = \log x, \quad x > 0.$$

Similarly the problem of finite summation is aided by the definition of the function

$$\underset{h}{L}(x, a) = \sum_{\nu=a}^{x-h} \frac{1}{\nu}, \quad a > 0, \quad \nu = a, a+h, a+2h, \ldots.$$

If $a$ is fixed this function is here defined only for the set of numbers $x = a+h$, $a+2h, \ldots$.

## 5. The summation of infinite series

The summation of convergent infinite series is a problem of importance in mathematics in spite of the fact that comparatively few important infinite series can be summed in closed form by direct methods. However, the methods of this chapter sometimes do permit of the summation of a series by direct method.

By definition, the sum of the convergent infinite series $\sum_{i=1}^{\infty} a_i$ is $\lim_{n\to\infty} \sum_{i=1}^{n-1} a_i$. Thus, the problem of evaluating the sum of the infinite series attacked directly is to evaluate the sum of the first $n-1$ terms and then the limit of this sum. Our discussion may help us in the first problem.

For example, $h = 1$,

$$\sum_{i=1}^{n-1} \frac{1}{(3i-2)(3i+1)(3i+4)} = -\frac{1}{3.2(3n-2)(3n+1)} + \frac{1}{24}.$$

Hence
$$\sum_{i=1}^{\infty} \frac{1}{(3i-2)(3i+1)(3i+4)} = \frac{1}{24}.$$

Summation by parts is frequently helpful in the approximate evaluation of the sum of a convergent infinite series. As an example we treat the series $\sum\limits_{i=1}^{\infty} 1/i^2$. In the summation by parts formula (3) let $u(i) = 1/i^2$, $\Delta v(i) = 1$, $v(i) = i$. Then

$$\begin{aligned}\sum_{i=1}^{\infty} \frac{1}{i^2} &= -1 + \sum_{i=1}^{\infty} \frac{2i+1}{i^2(i+1)} = -1 + 2\sum_{i=1}^{\infty} \frac{1}{i(i+1)} + \sum_{i=1}^{\infty} \frac{1}{i^2(i+1)} \\ &= -1 + 2\sum_{i=1}^{\infty} \frac{1}{i^2} - 2\sum_{i=1}^{\infty} \frac{1}{i^2(i+1)} + \sum_{i=1}^{\infty} \frac{1}{i^2(i+1)} \\ &= -1 + 2\sum_{i=1}^{\infty} \frac{1}{i^2} - \sum_{i=1}^{\infty} \frac{1}{i^2(i+1)}.\end{aligned}$$

Hence
$$\sum_{i=1}^{\infty} \frac{1}{i^2} = 1 + \sum_{i=1}^{\infty} \frac{1}{i^2(i+1)}.$$

The series in the right-hand member converges more rapidly than that in the left-hand member and consequently is better for approximate computing purposes. A like process can now be applied to $\sum\limits_{i=1}^{\infty} 1/i^2(i+1)$ and the computation made to depend upon a still more rapidly converging series. Details are left as an exercise.

The approximate summation of infinite series which we have just been discussing is also frequently facilitated by means of a transformation due to Euler. In order to prove Euler's theorem, we introduce two lemmas.

LEMMA I. *Hypotheses.*

(i) $\sum\limits_{k=0}^{\infty} A_k$ *converges, where* $A_k = \sum\limits_{n=0}^{\infty} a_{kn}$,

(ii) *Letting* $\sum\limits_{n=P}^{\infty} a_{kn} = {}_P r_k$, $\sum\limits_{k=0}^{\infty} {}_P r_k = R_P$ *converges for all values of* $P$,

(iii) $R_P \to 0$;

*Conclusion:* $\sum_{k=0}^{\infty} a_{kn}$ converges to a value which we call $B_n$ and

$$\sum_{n=0}^{\infty} B_n = \sum_{k=0}^{\infty} A_k.$$

*Proof:* $a_{kn} = {}_n r_k - {}_{n+1} r_k$ and consequently $\sum_{k=0}^{\infty} a_{kn} = B_n$ converges by hypothesis (ii). We write $A_k = A_k^P + {}_P r_k$. From which

$$\sum_{k=0}^{\infty} A_k = \sum_{k=0}^{\infty} A_k^P + \sum_{k=0}^{\infty} {}_P r_k = \sum_{n=0}^{P-1} B_n + R_P.$$

But $R_P \to 0$. The theorem follows.

LEMMA II. If

(i) $a_{np} \to 0$ for every fixed $p \geqslant 0$,

(ii) $\sum_{p=0}^{\infty} |a_{np}| < K$ for every value of $n > 0$,

(iii) $x_n \to 0$;

then $$\sigma_n = a_{n0} x_0 + a_{n1} x_1 + \ldots + a_{nn} x_n \to 0.$$

*Proof:* Let $\epsilon > 0$ be given, then if $M$ is large enough, when $n > M$, $|x_n| < \epsilon/2K$, and as a result

$$|\sigma_n| < |a_{n0} x_0 + a_{n1} x_1 + \ldots + a_{nM} x_M| + \tfrac{1}{2}\epsilon.$$

Now with $M$ held fast, choose $n$ so large that

$$|a_{nj} x_j| < \frac{\epsilon}{2M}$$

for $j = 0, \ldots, M$ simultaneously. Then $|\sigma_n| < \epsilon$ and the lemma is proved.

THEOREM II. *If* $$\sum_{n=0}^{\infty} a_n \tag{18}$$

*converges to $A$; then* $$\sum_{n=0}^{\infty} \frac{(1+E)^n a_0}{2^{n+1}}$$

*converges to $A$.*

*Proof:* Consider first $\dfrac{(1+E)^n a_k}{2^n}$, and let $a_{np}$ of Lemma II be $\dfrac{1}{2^n}\,{}_nC_p$ and $x_n = a_{k+n}$. We have the hypotheses of the lemma satisfied because

$$\frac{1}{2^n}\,{}_nC_p < \frac{1}{2^n} n^p \to 0$$

when $n \to \infty$ and

$$\sum_{p=0}^{n} \frac{1}{2^n} {}_nC_p = \frac{1}{2^n}(1+1)^n = 1.$$

Of course $a_n \to 0$.

Hence
$$\lim_{n\to\infty} \frac{1}{2^n}(1+E)^n a_k = 0. \tag{19}$$

Next, let
$$a_{kn} = \left[\frac{(1+E)^n}{2^n} a_k - \frac{(1+E)^{n+1}}{2^{n+1}} a_k\right].$$

Then, by (19),

$$\sum_{n=0}^{\infty} a_{kn} = \frac{1}{2^0}(1+E)^0 a_k = a_k.$$

We thus obtain an infinite series for each term of (18). We wish to apply Lemma I to the resulting array.

On account of the relation

$$(1+E)^{n+1}a_k = (1+E)^n a_k + (1+E)^n a_{k+1}$$

we can write
$$\sum_{k=0}^{\infty} \left(\frac{(1+E)^n}{2^n} a_k - \frac{(1+E)^{n+1}}{2^{n+1}} a_k\right)$$

in the form

$$\sum_{k=0}^{\infty} \frac{1}{2^{n+1}}[(1+E)^n a_k - (1+E)^n a_{k+1}]$$
$$= \frac{1}{2^{n+1}}\left[(1+E)^n a_0 - \lim_{k\to\infty}(1+E)^n a_{k+1}\right].$$

Since $a_k \to 0$, and since there are a fixed number of terms in $(1+E)^n a_{k+1}$, we see that $(1+E)^n a_{k+1} \to 0$ when $k \to \infty$. Consequently,

$$\sum_{k=0}^{\infty} a_{kn} = \frac{(1+E)^n}{2^{n+1}} a_0.$$

In order for the theorem to be complete, it is now only necessary to show that $R_P$ of Lemma I approaches zero. Using the notation of that theorem,

$${}_P r_k = \frac{(1+E)^P a_k}{2^P},$$

and consequently,

$$R_P = \frac{1}{2^P} \sum_{k=0}^{\infty} (1+E)^P a_k.$$

Now, let
$$a_k + a_{k+1} + \ldots = r_k.$$
Write $(1+E)^P a_k$ in expanded form and we have
$$R_P = \frac{1}{2^P} \sum_{i=0}^{P} {}_P C_i \, r_i \to 0$$
by Lemma II, since $r_P \to 0$.

THEOREM III. *Hypothesis:*
$$\sum_{n=0}^{\infty} a_n$$
*converges to* $A$.

*Conclusion:*
$$\frac{1}{c} \sum_{n=0}^{\infty} \left[\frac{(c-1)+E}{c}\right]^n a_0, \quad c \geqslant 1,$$
*converges to* $A$ *also.*

This theorem, which reduces to the Euler theorem in case $c = 2$, can be proved with but slight modifications of the proof given for that theorem. Proof is consequently omitted. The transformation will be called the extended Euler transformation.

A series
$$\sum_{n=1}^{\infty} a_n \tag{20}$$
can be written
$$\sum_{n=1}^{\infty} \Delta s_{n-1}, \quad \Delta s_{n-1} = s_n - s_{n-1}, \tag{21}$$
if
$$s_n = \sum_{n=1}^{n} a_n, \quad n \geqslant 1, \quad s_0 = 0.$$
Substitute for (20) the series
$$\sum_{n=1}^{\infty} (\Delta s_{n-1} + \Delta b_{n-1}), \tag{22}$$
where $b_n \to 0$ and $b_0 = 0$. The sum of the series will not be changed. (The replacement of (20) by (22) is called a *modification of* (20).) Now it is frequently possible to choose $b_n$ so that (22) converges more rapidly than (20).

For example, consider
$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n}. \tag{23}$$

Let $b_n = (-1)^n \frac{1}{2n+1}$. Then

$$\sum_{n=2}^{\infty} (\Delta s_{n-1} + \Delta b_{n-1})$$
$$= \sum_{n=2}^{\infty} \left[ (-1)^{n-1}\frac{1}{n} + (-1)^n \frac{1}{2n+1} - (-1)^{n-1}\frac{1}{2n-1} \right]$$
$$= \sum_{n=2}^{\infty} (-1)^n \frac{1}{n(2n-1)(2n+1)}.$$

This is a series which converges much more rapidly than (23).

In general we consider a series

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{P(n)},$$

where $P(n)$ is a polynomial of degree $\rho \geqslant 1$. Let

$$b_n = (-1)^n \frac{u(n)}{v(n)},$$

where $u(n)$ and $v(n)$ are to be polynomials. Then

$$\Delta s_{n-1} + \Delta b_{n-1}$$
$$= (-1)^{n-1} \frac{v(n)v(n-1) - P(n)\{u(n)v(n-1) + u(n-1)v(n)\}}{P(n)v(n)v(n-1)}. \tag{24}$$

We now endeavour to choose the polynomials $u(n)$ and $v(n)$ so that the degree of the numerator in (24) is as much less than the degree of the denominator as possible. Let the degree of $v(n)$ be $\sigma$ and that of $u(n)$ be $\tau$ and let $\sigma = \tau + \rho$. Terms in the numerator of degree $2\sigma$ formally appear. However, there are $\tau+1$ coefficients in $u(n)$ and $\tau+\rho+1$ coefficients in $v(n)$. We hope to be able so to choose these coefficients as to reduce the degree of the numerator from the apparent $2\tau+2\rho$ to $\rho-2$. The degree of the denominator will remain $\rho+2\sigma$. But

$$\sigma = \tau + \rho \geqslant \rho.$$

Consequently, the degree of the denominator is at least $3\rho$.

## EXERCISES

1. Work out, $h = h$,

(a) $\sum x^3$, (b) $\sum (2x^4+3x^3-1)$,

(c) $\sum \frac{1}{x(x+h)}$, (d) $\sum \frac{1}{(x+5h)(x+3h)}$,

(e) $\sum x^2 \sin x$.

2. Evaluate, $h = 1$,

(a) $\sum_{x=1}^{999} x^5$, (b) $\sum_{x=1}^{n} Ma^x$,

(c) $\sum_{x=1}^{100} \sin 2x$, (d) $\sum_{x=1}^{100} 10^x \cos 3x$,

(e) $\sum_{x=1}^{\infty} \frac{1}{x(x-1)(x-2)}$, (f) $\sum_{x=1}^{\infty} e^{-x} \sin 3x$,

(g) $\sum_{x=1}^{n} \sinh x$, (h) $\sum_{x=1}^{n} \cosh x$.

3. Prove $$\sum u(x) = -\sum_{\nu=0}^{\infty} u(x+\nu h)+c$$
in case the infinite series is convergent.

4. Derive the summation by parts formula
$$\sum \phi(x+h)\Delta^2\psi(x) = \phi(x)\Delta\psi(x)-\psi(x)\Delta\phi(x) + \sum \psi(x+h)\Delta^2\phi(x).$$

5. Generalize the formula in Exercise 4.

6. Use summation by parts to replace $\sum_{i=1}^{\infty} 1/i^2(i+1)$ by a constant plus a more rapidly convergent series. Then, noting that $\frac{1}{6}\pi^2 = \sum_{i=1}^{\infty} 1/i^2$, use the results that you have obtained and the results of the text to calculate $\pi^2$ to a few decimal figures.

7. By the use of summation by parts obtain from the formula
$$\sum_{i=1}^{\infty} (-1)^{i-1}\frac{1}{i} = \log 2$$
a more rapidly converging series for $\log 2$.

8. Apply the Euler transformation to $\sum_{i=1}^{\infty} (-1)^i \, 1/i$ and then calculate $\log 2$, correct to four decimal places.

9. Apply the Euler transformation to $\sum_{i=1}^{\infty} (-1)^{i-1} \, 1/i^2$ and then calculate its sum to a few decimal places.

10. If $E_C$ denotes the extended Euler transformation with constant $C$, prove that

$$E_{C_1} E_{C_2} = E_{C_1 C_2}.$$

11. Give in detail a proof of Theorem III.

12. Modify the series

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{2n-1} = \tfrac{1}{4}\pi$$

by the method of p. 22, and use the modified series to calculate $\pi$ to seven decimal places.

13. Modify the series $$\sum_{n=1}^{\infty} \frac{1}{n^2} = \tfrac{1}{6}\pi^2$$

by the method of p. 22, and use the modified series to calculate $\pi^2$ to several decimal places.

## SUGGESTED REFERENCES

BRADSHAW, J. W.: 'Modified Series', *Amer. Math. Monthly*, **46**, 286; 'The Modification of an Infinite Product', *Quart. J. of Math.* (Oxford), **12**, 216.

SHOHAT, J. A.: 'On a Certain Transformation of Infinite Series', *Amer. Math. Monthly*, **40**, 226.

FORT, T.: *Infinite Series*, chaps. VII and VIII.

# III

# THE BERNOULLI AND EULER POLYNOMIALS AND NUMBERS

## 1. Definitions and some fundamental formulae

THE process of summation is facilitated by the introduction of certain polynomials known as Bernoulli polynomials. These not only play an important role in the theory of finite differences but serve as a natural introduction to the Bernoulli numbers which occur frequently in important parts of mathematics.

For the time being we limit the independent variable to the succession of values 0, 1, 2, 3,... .

Let

$$\phi_n(x) = \sum_{x=0}^{x-1} nx^{n-1}, \quad n > 0; \qquad \phi_n(0) = 0; \tag{1}$$

$$\phi_0(x) \equiv 0. \tag{1'}$$

Now by Newton's formula† we can write

$$nx^{n-1} = nx^{(n-1)} + a_1 x^{(n-2)} + \ldots + a_{n-1} x^{(1)},$$

where $a_1,\ldots, a_{n-1}$ are constants. Apply the operator $\sum_{x=0}^{x-1}$ to both members of this and we have

$$\phi_n(x) = x^{(n)} + \frac{a_1}{n-1} x^{(n-1)} + \ldots + \frac{a_{n-2}}{2} x^{(2)}, \quad n > 1. \tag{2}$$

This is a polynomial of the $n$th degree which we write

$$\begin{aligned} \phi_n(x) &= B_0^{(n)} x^n + {}_nC_1\, B_1^{(n)} x^{n-1} + \ldots + {}_nC_{n-1}\, B_{n-1}^{(n)} x \\ &= \sum_{\nu=0}^{n-1} {}_nC_\nu\, B_\nu^{(n)} x^{n-\nu}, \end{aligned} \tag{3}$$

where the $B$'s are independent of $x$.

From (2) it is immediate that

$$\phi_n(0) = \phi_n(1) = 0, \quad n > 1. \tag{3'}$$

From (1)

$$\Delta\phi_n(x) = \phi_n(x+1) - \phi_n(x) = nx^{n-1}. \tag{4}$$

The polynomials $\phi_n(x)$, although till now defined for $x = 0, 1, 2,\ldots$ only, are defined by the formula in the right-hand member of (3) for any value of $x$ and henceforth will be

† Chap. I, § 7.

considered as algebraic polynomials. Since relation (4) holds for all positive integral values of $x$, it is an identity.

Differentiate both sides of (4), and denote differentiation by an accent. We get

$$\frac{1}{n}\Delta\phi_n'(x) = (n-1)x^{n-2}. \tag{5}$$

Sum this from 0 to $x-1$ over integral values. We get

$$\begin{aligned}\frac{1}{n}[\phi_n'(x)-\phi_n'(0)] &= \sum_{x=0}^{x-1}(n-1)x^{n-2} \\ &= \phi_{n-1}(x) = \sum_{\nu=0}^{n-2} {}_{n-1}C_\nu\, B_\nu^{(n-1)}x^{n-1-\nu}.\end{aligned} \tag{6}$$

This again is an algebraic identity in $x$. However, from (3), by differentiation,

$$\frac{1}{n}[\phi_n'(x)-\phi_n(0)] = \sum_{\nu=0}^{n-2} {}_{n-1}C_\nu\, B_\nu^{(n)}x^{n-1-\nu}.$$

The right-hand member of this is identical with the right-hand member of (6). In other words, $B_0^{(n)}$, $B_1^{(n)}$, $B_2^{(n)}$,... are independent of $n$. They will be denoted by $B_0$, $B_1$, $B_2$,... .

To the polynomial written in (1) it is frequently convenient to add $B_n$. Whereupon we write

$$B_n(x) \equiv \phi_n(x)+B_n = \sum_{\nu=0}^{n} {}_nC_\nu\, B_\nu\, x^{n-\nu} = (x+B)^n, \tag{7}$$

where in the binomial expansion exponents are to be written to $x$ and subscripts to $B$.

*The polynomial $B_n(x)$, $n = 1, 2,...,$ will be called the Bernoulli polynomial of degree $n$. The numbers $B_0$, $B_1$, $B_2$,... will be called the Bernoulli*† *numbers.*

† Definitions of the Bernoulli numbers and polynomials which yield closely related but slightly different results have been given from time to time. However, definitions equivalent to those which are given here have certain elements of convenience. In particular, numbers $\beta_k$ determined by

$$\beta_k = (-1)^{k-1}B_{2k}, \quad k > 0,$$

are frequently called Bernoulli numbers and the polynomials defined by

$$\chi_n(x) = \sum_{x=0}^{x-1} \frac{x^{n-1}}{(n-1)!}$$

are frequently called Bernoulli polynomials.

From (3) and (3′) we have

$$\phi_n(1) = B_0 + {}_nC_1 B_1 + {}_nC_2 B_2 + \ldots + {}_nC_{n-1} B_{n-1} = 0. \tag{8}$$

This is a formula from which $B_{n-1}$ can be calculated if $B_0, \ldots, B_{n-2}$ are known. Calculations, however, are increasingly laborious and the necessary labour soon becomes prohibitive. From (2) $B_0 = 1$. Starting with $B_0 = 1$, we find $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_3 = 0$, $B_4 = -\frac{1}{30}$, $B_5 = 0$, $B_6 = \frac{1}{42}, \ldots$.

## 2. Further fundamental formulae

From (1), (6), and (7)

$$\frac{dB_n(x)}{dx} = nB_{n-1}(x), \quad n \geqslant 1; \tag{9}$$

$$\frac{d^pB_n(x)}{dx^p} = n(n-1)\ldots(n-p+1)B_{n-p}(x), \quad p \geqslant 1,\ n \geqslant 1. \tag{9′}$$

From (9)

$$\int\limits_a^b B_n(x)\,dx = \frac{1}{n+1}[B_{n+1}(b) - B_{n+1}(a)], \quad n \geqslant 1. \tag{10}$$

In particular, since from (3′), $B_n(0) = B_n(1)$, $n > 1$,

$$\int\limits_0^1 B_n(x)\,dx = 0, \quad n \geqslant 1. \tag{10′}$$

From (1) we obtain the important relation

$$1^n + 2^n + \ldots + (x-1)^n = \sum_{x=1}^{x-1} x^n = \frac{B_{n+1}(x) - B_n}{n+1}. \tag{11}$$

If we expand $B_n(x+h)$ by Taylor's formula and apply (9′) we obtain

$$B_n(x+h) = \sum_{\nu=0}^{n} {}_nC_\nu B_{n-\nu} h^\nu.$$

If in this we let $h = 1$ and subtract $B_n(x)$ we get

$$\begin{aligned} \Delta B_n(x) &= B_n(x+1) - B_n(x) \\ &= \sum_{\nu=0}^{n} {}_nC_\nu B_{n-\nu}(x) - B_n(x) = \sum_{\nu=0}^{n-1} {}_nC_\nu B_\nu(x). \end{aligned} \tag{12}$$

Compare this with (4), and we have, since $\Delta\phi_n(x) = \Delta B_n(x)$,

$$\sum_{\nu=0}^{n-1} {}_nC_\nu B_\nu(x) = nx^{n-1}. \tag{13}$$

This is a formula from which $B_{n-1}(x)$ can be calculated if $B_0(x)$, $B_1(x),\ldots,$ $B_{n-2}(x)$ are known. Since $B_0 = 1$, from (1′) and (7) $B_0(x) = 1$. We readily find

$$\begin{aligned}
B_0(x) &= 1,\\
B_1(x) &= x-\tfrac{1}{2},\\
B_2(x) &= x^2-x+\tfrac{1}{6},\\
B_3(x) &= x^3-\tfrac{3}{2}x^2+\tfrac{1}{2}x,\\
B_4(x) &= x^4-2x^3+x^2-\tfrac{1}{30},\\
B_5(x) &= x^5-\tfrac{5}{2}x^4+\tfrac{5}{3}x^3-\tfrac{1}{6}x,\\
B_6(x) &= x^6-3x^5+\tfrac{5}{2}x^4-\tfrac{1}{2}x^2+\tfrac{1}{42},\\
B_7(x) &= x^7-\tfrac{7}{2}x^6+\tfrac{7}{2}x^5-\tfrac{7}{6}x^3+\tfrac{1}{6}x,\\
B_8(x) &= x^8-4x^7+\tfrac{14}{3}x^6-\tfrac{7}{3}x^4+\tfrac{2}{3}x^2-\tfrac{1}{30},\\
B_9(x) &= x^9-\tfrac{9}{2}x^8+6x^7-\tfrac{21}{5}x^5+2x^3-\tfrac{3}{10}x,\\
B_{10}(x) &= x^{10}-5x^9+\tfrac{15}{2}x^8-7x^6+5x^4-\tfrac{3}{2}x^2+\tfrac{5}{66}.
\end{aligned}$$

Of course, these polynomials can be found in other ways; for example, by writing $nx^{n-1}$ as a sum of factorials and summing as indicated in § 1, or by calculating the Bernoulli numbers and substituting in $(x+B)^n$ or as indicated in the following section.

## 3. Integral formulae for calculating $B_n$ and $B_n(x)$

Integration formulae can be developed for the successive calculation of $B_n$ as follows: Successive integration by parts applied to the left-hand number yields the following formula,

$$\int_0^1 B_\mu(x)B_\nu(x)\,dx = \frac{(-1)^{\mu-1}(\mu!)(\nu!)}{(\mu+\nu)!}B_{\mu+\nu}, \quad \mu,\nu > 0. \tag{14}$$

Interesting special cases of this formula are

$$\int_0^1 B_{2\nu}^2(x)\,dx = \frac{-\{(2\nu)!\}^2}{(4\nu)!}B_{4\nu}, \quad \nu > 0; \tag{15}$$

$$\int_0^1 B_{2\nu+1}^2(x)\,dx = \frac{(2\nu+1)!}{(4\nu+2)!}B_{4\nu+2}, \quad \nu > 0; \tag{16}$$

$$\int_0^1 B_{2\nu}(x)B_2(x)\,dx = -\frac{1}{(\nu+1)(2\nu+1)}B_{2\nu+2}, \quad \nu > 0; \tag{17}$$

$$\int_0^1 B_{2\nu-1}(x)B_1(x)\,dx = \frac{1}{(2\nu)!}B_{2\nu}, \quad \nu > 0. \tag{18}$$

If in (14) we interchange $\mu$ and $\nu$ we have

$$(-1)^{\mu-1}B_{\mu+\nu} = (-1)^{\nu-1}B_{\nu+\mu}.$$

Hence, if $n$ is odd $B_n = 0$, that is

$$B_{2\nu+1} = 0, \quad \nu > 0. \tag{19}$$

As a matter of fact on account of (10′)

$$\int_0^1 B_{2\nu-1}(x)\,dx = 0,$$

and since by (10′) and (14) and (19)

$$\int_0^1 [-x+\tfrac{1}{6}]B_{2\nu}(x)\,dx = \int_0^1 \{-B_1(x)B_{2\nu}(x)\}\,dx - \tfrac{1}{3}\int_0^1 B_{2\nu}(x)\,dx = 0,$$

for purposes of calculation we can replace (17) and (18) by

$$\int_0^1 x^2 B_{2\nu}(x)\,dx = -\frac{1}{(\nu+1)(2\nu+1)}B_{2\nu+2}, \quad \nu > 0, \tag{20}$$

and

$$\int_0^1 x B_{2\nu-1}(x)\,dx = \frac{1}{2\nu}B_{2\nu}, \quad \nu > 0. \tag{21}$$

From formulae (7) and (9)

$$B_n(x) - B_n = n\int_0^x B_{n-1}(x)\,dx. \tag{22}$$

If we have the Bernoulli numbers this gives a means of successive calculation of $B_n(x)$ from $B_1(x) = x-\frac{1}{2}$. The labour is great for a large value of $n$.

## 4. A difference relationship

We have seen that $B_n(x) = (x+B)^n$, where in the binomial expansion exponents are to be written to $x$ and subscripts to $B$. Now take a parameter $h$, then

$$B_n(x+h) = (x+h+B)^n = (x+B+h)^n. \tag{23}$$

Let $\psi(x)$ be a polynomial of the $m$th degree. Form

$$\psi(x+B+1)-\psi(x+B).$$

Replace in the expression thus formed $(x+B+1)^k$ and $(x+B)^k$ by $B_k(x+1)$ and $B_k(x)$ respectively, $k = 0,\ldots, m$.

Apply (4) and we have

$$\psi(x+B+1)-\psi(x+B) = \psi'(x). \tag{24}$$

In other words, $\psi(x+B)$ is a polynomial solution of the equation

$$\Delta y = \psi'(x),$$

where the right-hand member is any given polynomial.

Equation (24) can be written

$$\psi\{B(x)+1\}-\psi\{B(x)\} = \psi'(x).$$

If in this we let $\psi(x) = x^n$ we have again formula (13), which we write as

$$[B(x)+1]^n-B(x) = nx^{n-1}. \tag{25}$$

If we let $x = 0$ we get a second time relation (8), which we write

$$(B+1)^n-B_n = 0, \quad n > 1. \tag{26}$$

## 5. The Euler-Maclaurin summation formula for polynomials

As previously, let $\psi(x)$ be a polynomial of the $m$th degree and in the expansion of $\psi(x+B)$ apply subscripts and not exponents to $B$. Then

$$\frac{d}{dx}\psi(x+B) = \psi'(x+B).$$

Hence the polynomial $\psi(x+B)$ behaves exactly as $\psi(x+h)$ with exponents applied to $h$ not only as regards algebraic operations but also as regards differentiation. We consequently can use Taylor's formula. We have

$$\psi(x+B) = \sum_{\nu=0}^{m} \frac{B_\nu}{\nu!}\psi^{(\nu)}(x). \tag{27}$$

Similarly, with arbitrary but fixed $q$,

$$\psi(x+q+B) = \psi(x+B+q) = \psi\{x+(B+q)\}$$

$$= \psi\{x+B(q)\} = \sum_{\nu=0}^{m} \frac{B_\nu(q)}{\nu!}\psi^{(\nu)}(x). \tag{28}$$

Taking the first difference of (28) and applying (24), we get

$$\psi'(x+q) = \sum_{\nu=0}^{m-1} \frac{B_\nu(q)}{\nu!} \Delta\psi^{(\nu)}(x).$$

Replace $\psi(x)$ by $\int_0^x \chi(x)\,dx$, where $\chi(x)$ is a polynomial of degree $m-1$. Since

$$\Delta\psi^{(0)}(x) = \Delta\psi(x) = \psi(x+1)-\psi(x) = \int_x^{x+1} \chi(x)\,dx,$$

we obtain

$$\chi(x+q) = \int_x^{x+1} \chi(x)\,dx + \sum_{\nu=1}^{m-1} \frac{B_\nu(q)}{\nu!} \Delta\chi^{(\nu-1)}(x). \tag{29}$$

This is a polynomial identity. It is called *the Euler-Maclaurin sum formula* and will be discussed further later.

## 6. Symmetry property

We know that

$$\Delta B_{2k+1}(x) = (2k+1)x^{2k}.$$

Replace $x$ by $-x$ and we obtain

$$B_{2k+1}(1-x) - B_{2k+1}(-x) = (2k+1)x^{2k}.$$

Let $$-B_{2k+1}(1-x) = F(x).$$

Then $$\Delta F(x) = (2k+1)x^{2k}.$$

Now, if two polynomials have the same difference they differ at most by a constant. We know

$$B_{2k+1}(0) = B_{2k+1} = B_{2k+1}(1) = -F(0) = -F(1).$$

Hence $$F(x) = B_{2k+1}(x) - 2B_{2k+1}.$$

That is, $$-B_{2k+1}(1-x) = B_{2k+1}(x) - 2B_{2k+1}. \tag{30}$$

Differentiate both sides of (30) and apply (9). We get

$$B_{2k}(1-x) = B_{2k}(x). \tag{31}$$

Differentiate (31) and apply (9). We arrive at

$$B_{2k-1}(1-x) = -B_{2k-1}(x). \tag{32}$$

Equations (31) and (32) can be combined into

$$B_n(1-x) = (-1)^n B_n(x). \tag{33}$$

This is a fundamental relation in the theory of Bernoulli polynomials.

In (33) let $x = 0$ and we have

$$B_{2k+1} = -B_{2k+1}, \quad k > 0.$$

Hence as we already know

$$B_{2k+1} = 0, \quad k > 0.$$

## 7. The polynomials $\phi_n(x) = B_n(x) - B_n$ in the interval $(0, 1)$

Put $x = \frac{1}{2}$ in (33) and we see that

$$B_{2k+1}(\tfrac{1}{2}) = 0.$$

Moreover, since $B_{2k+1} = 0$,

$$B_{2k+1}(0) = B_{2k+1}(1) = 0, \quad k > 0.$$

THEOREM. *There is no zero of* $\phi_{2k+1}(x) = B_{2k+1}(x)$ *within the interval* $(0, 1)$ *other than* $\frac{1}{2}$.

We prove this by contradiction. Assume zeros, $a_1$ and $a_2$, such that

$$0 < a_1 < a_2 < 1.$$

Here either $a_1$ or $a_2$ might be $\frac{1}{2}$. By Rolle's theorem, $B'_{2k+1}(x)$ has at least three zeros, $b_1$, $b_2$, and $b_3$, such that

$$0 < b_1 < a_1 < b_2 < a_2 < b_3 < 1.$$

Hence $B''_{2k+1}(x)$ has at least two zeros within the interval $(0, 1)$. But

$$B''_{2k+1}(x) = (2k+1)(2k)B_{2k-1}(x).$$

Hence $B_{2k-1}(x)$ has at least two zeros within the interval $(0, 1)$. By successive repetitions, $B_3(x)$ must have at least two zeros in this interval, but this is contrary to the fact. An explicit form for $B_3(x)$ has been given on p. 29.

THEOREM. $\phi_{2k}(x) = B_{2k}(x) - B_{2k}$ *retains the same sign over the interval* $(0, 1)$.

If $\phi_{2k}(x)$ has a zero on this interval, since $\phi_{2k}(0) = \phi_{2k}(1) = 0$, $\phi'_{2k}(x)$ will have at least two. But $\phi'_{2k}(x) = 2kB_{2k-1}(x)$ and we have just seen that this has but one zero on this interval.

THEOREM. $\phi_{2k+1}(x)$ *has a simple zero when* $x = \frac{1}{2}$.

Proof is as follows. Suppose

$$\phi'_{2k+1}(\tfrac{1}{2}) = (2k+1)[\phi_{2k}(\tfrac{1}{2}) + B_{2k}] = 0.$$

Since $\phi_{2k+1}(0) = \phi_{2k+1}(\frac{1}{2}) = 0$, by Rolle's theorem, $\phi'_{2k+1}(x)$ has a zero within the interval $(0, \frac{1}{2})$. Hence, again by Rolle's theorem, $\phi''_{2k+1}(x)$ has a zero between 0 and $\frac{1}{2}$, but

$$\phi''_{2k+1}(x) = (2k+1)(2k)\phi_{2k-1}(x)$$

and $\phi_{2k-1}(x)$ has no zero on this interval.

THEOREM. $$(-1)^{k+1}B_{2k} > 0.$$

When $k > 1$ this follows from (15) and (16). We know that

$$B_2 = \tfrac{1}{6}.$$

THEOREM. *The function $\phi_{2k}(x)$ is of the same sign as $B_{2k-2}$ over the interval $0 < x < 1$ and $\phi_{2k+1}(x)$ $(k > 0)$ is of the same sign as $B_{2k}$ over the interval $0 < x < \frac{1}{2}$.*

This theorem follows from the fact that $\phi_n(0) = 0$ and that the polynomials in question do not vanish on the intervals mentioned and consequently must have over these intervals the signs of the coefficients of their terms of lowest degree.

## 8. Multiplication theorem

Consider the equation

$$f\left(x+\frac{1}{m}\right)-f(x) = nx^{n-1}.$$

Trial, with reference to equation (4), shows that each of the following is a polynomial solution:

$$\sum_{\nu=0}^{m-1} B_n\left(x+\frac{\nu}{m}\right) \quad \text{and} \quad m^{1-n}B_n(mx).$$

These then differ at most by a constant. We write

$$\sum_{\nu=0}^{m-1} B_n\left(x+\frac{\nu}{m}\right) = m^{1-n}B_n(mx)+C.$$

Integrate both sides of this equation from $x$ to $x+1$ and apply formula (10) and then (4) to the left-hand member. We get

$$\frac{1}{m^n}[(mx)^n+(mx+1)^n+...+(mx+m-1)^n]$$
$$= \frac{B_{n+1}(mx+m)-B_{n+1}(mx)}{(n+1)m^n}+C.$$

Now apply formula (11) to the left-hand member of this equation and we find $C = 0$. We then write

$$\sum_{\nu=0}^{m-1} B_n\left(x+\frac{\nu}{m}\right) = m^{1-n}B_n(mx). \tag{34}$$

This has been called a 'multiplication theorem' for the Bernoulli polynomials. From it there can be obtained some formulae which are useful in plotting the graph of $B_n(x)$. Let $x = 0$ in (34), and we obtain the following relation,

$$\sum_{\nu=1}^{m-1} B_n\left(\frac{\nu}{m}\right) = -\left(1-\frac{1}{m^{n-1}}\right)B_n. \tag{35}$$

This coupled with (33) yields

$$B_n\left(\frac{1}{2}\right) = -\left(1-\frac{1}{2^{n-1}}\right)B_n, \quad n = 1, 2,...,$$

$$B_n\left(\frac{1}{3}\right) = B_n\left(\frac{2}{3}\right) = -\left(1-\frac{1}{3^{n-1}}\right)\frac{B_n}{2}, \quad n \text{ even},$$

$$B_n\left(\frac{1}{4}\right) = B_n\left(\frac{3}{4}\right) = -\left(1-\frac{1}{2^{n-1}}\right)\frac{B_n}{2^n}, \quad n \text{ even},$$

$$B_n\left(\frac{1}{6}\right) = B_n\left(\frac{5}{6}\right) = \left(1-\frac{1}{2^{n-1}}\right)\left(1-\frac{1}{3^{n-1}}\right)\frac{B_n}{2}, \quad n \text{ even}. \tag{36}$$

## 9. Fourier development of the Bernoulli polynomials over (0, 1)

Consider first $B_{2k}(x)$. We know that there exists a Fourier expansion for this polynomial valid over the interval (0, 1). We write

$$B_{2k}(x) = \tfrac{1}{2}a_0^{(2k)}+a_1^{(2k)}\cos 2k\pi x+b_1^{(2k)}\sin 2k\pi x+...,$$

where

$$a_n^{(2k)} = 2\int_0^1 B_{2k}(x)\cos 2n\pi x\, dx, \qquad b_n^{(2k)} = 2\int_0^1 B_{2k}(x)\sin 2n\pi x\, dx.$$

Here $b_n^{(2k)} = 0$ since

$$B_{2k}(x) = B_{2k}(1-x).$$

Next,
$$a_0^{(2k)} = 2\int_0^1 B_{2k}(x)\, dx.$$

Utilize (10) and we have

$$a_0^{(2k)} = \frac{2}{2k+1}[B_{2k+1}(1)-B_{2k+1}(0)] = 0.$$

In order to determine $a_n^{(2k)}$, $n > 0$, we proceed as follows: Use the integration by parts formula

$$\int \phi\psi''\,dx = \phi\psi'-\phi'\psi+\int \psi\phi''\,dx, \tag{37}$$

with $\phi = B_{2k}(x), \qquad \psi'' = \cos 2n\pi x,$

then $\phi'' = 2k(2k-1)B_{2k-2}(x), \qquad \psi = -\dfrac{\cos 2n\pi x}{(2n\pi)^2}.$

We first find

$$a_n^{(2k)} = \frac{-2k(2k-1)}{(2n\pi)^2}a_n^{(2k-2)}, \quad n > 0.$$

Now we know that

$$B_2(x) = x^2-x+\tfrac{1}{6}.$$

We then find $a_n^{(2)} = \dfrac{2.2}{(2n\pi)^2}.$

So that, finally,

$$a_n^{(2k)} = \frac{(-1)^{k-1}2.2k(2k-1)...2}{(2n\pi)^{2k}}, \quad n > 0,\ k > 0. \tag{38}$$

This Fourier expansion is valid when $x = 0$ and $x = 1$ as well as over the interval $0 < x < 1$, since $B_{2k}(0) = B_{2k}(1)$.

Next consider $B_{2k-1}(x)$ and use an analogous notation to what has just been used. It is immediate that

$$\tfrac{1}{2}a_0^{(2k-1)} = B_{2k-1}, \qquad a_n^{(2k-1)} = 0, \quad n > 0.$$

$$b_n^{(2k-1)} = 2\int_0^1 B_{2k-1}(x)\sin 2n\pi x\,dx.$$

Proceeding as before we find

$$b_n^{(1)} = -\frac{1}{n\pi},$$

$$b_n^{(2k-1)} = \frac{-(2k-1)(2k-2)}{(2n\pi)^2}b_n^{(2k-3)}, \quad k > 1,$$

whence

$$b_n^{(2k-1)} = \frac{(-1)^k.2.(2k-1)(2k-2)...2.1}{(2n\pi)^{2k-1}}, \quad n > 0,\ k > 0. \tag{39}$$

The expansion thus obtained is again valid when $x = 0$, and $x = 1$ if $2k-1 > 1$. We write

$$B_{2k}(x) = \frac{(-1)^{k-1}.2.\{(2k)!\}}{(2\pi)^{2k}}\left[\cos 2\pi x + \frac{\cos 4\pi x}{2^{2k}} + \frac{\cos 6\pi x}{3^{2k}} + \ldots\right], \tag{40}$$

$$B_{2k-1}(x) = \frac{(-1)^{k}.2.\{(2k-1)!\}}{(2\pi)^{2k-1}}\left[\sin 2\pi x + \frac{\sin 4\pi x}{2^{2k-1}} + \frac{\sin 6\pi x}{3^{2k-1}} + \ldots\right]. \tag{41}$$

If in (40) we let $x = 0$, we get the remarkable relation

$$\begin{aligned} B_{2k} &= \frac{(-1)^{k-1}.2.\{(2k)!\}}{(2\pi)^{2k}}\left[1 + \frac{1}{2^{2k}} + \frac{1}{3^{2k}} + \ldots\right] \\ &= \frac{(-1)^{k-1}.2.\{(2k)!\}}{(2\pi)^{2k}}\zeta(2k), \end{aligned} \tag{42}$$

where $\zeta(2k)$ is the celebrated $\zeta$-function. If we let $k = 1, 2$ in (42) we get

$$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \ldots, \qquad \frac{\pi^4}{90} = 1 + \frac{1}{2^4} + \frac{1}{3^4} + \frac{1}{4^4} + \ldots.$$

More important, however, are the relations (43) and (44) below. We observe that

$$\begin{aligned} 1 < \zeta(2k) &= 1 + \left(\frac{1}{2^{2k}} + \frac{1}{3^{2k}}\right) + \left(\frac{1}{4^{2k}} + \frac{1}{5^{2k}} + \frac{1}{6^{2k}} + \frac{1}{7^{2k}}\right) + \ldots \\ &< 1 + \frac{1}{2^{2k-1}} + \frac{1}{2^{4k-2}} + \frac{1}{2^{6k-3}} + \ldots = \frac{1}{1-1/2^{2k-1}}. \end{aligned}$$

This right-hand member approaches 1 as $k \to \infty$. It is immediate that

$$\frac{2\{(2k)!\}}{(2\pi)^{2k}} < |B_{2k}| < \frac{2\{(2k!)\}}{(2\pi)^{2k}}\frac{1}{1-1/2^{2k-1}}. \tag{43}$$

We next conclude from consideration of (40) that

$$|B_{2k}(x)| < |B_{2k}|, \quad 0 < x < 1.$$

Similarly from (41) we conclude

$$|B_{2k-1}(x)| < \frac{2\{(2k-1)!\}}{(2\pi)^{2k-1}}\frac{1}{1-1/2^{2k-2}}. \tag{44}$$

## 10. Theorem of Jacobi

Some of the properties of Bernoulli polynomials can be established by means of the following theorem known as the Theorem of Jacobi. Notable among these is the symmetry property of § 6.

THEOREM. *If* $z = x-x^2$ *then*

$$(-1)^k B_{2k}(x) = \sum_{\nu=0}^{k} d_{\nu k}\, z^{k-\nu} \tag{45}$$

*and* $$(-1)^k B_{2k-1}(x) = (1-2x)\sum_{\nu=0}^{k-2} \delta_{\nu k}\, z^{k-\nu-1}, \quad k > 1, \tag{46}$$

*where the* $d_{\nu k}$ *and* $\delta_{\nu k}$ *are constants,* $\left.\begin{matrix} d_{\nu k} \\ \delta_{\nu k}\end{matrix}\right\} \geqslant 0.$

We shall prove formula (45) by showing that $d_{\nu k}$ can be determined to satisfy (45). Formula (46) will be left to the reader.

Differentiate (45) twice and utilize (9). We get

$$(-1)^k (2k)^{(2)} B_{2k-2}(x) = \sum_{\nu=0}^{k} d_{\nu k}[(k-\nu)^{(2)} z^{k-\nu-2} - (2k-2\nu)^{(2)} z^{k-\nu-1}]$$
$$= \sum_{\nu=1}^{k-1} [(k-\nu+1)^{(2)} d_{(\nu-1)k} - (2k-2\nu)^{(2)} d_{\nu k}] z^{k-\nu-1} - (2k)^{(2)} d_{0k}\, z^{k-1}.$$

Now if we write (45) with $k$ replaced by $(k-1)$ and equate coefficients we obtain

$$(2k-2\nu)^{(2)} d_{\nu k} = (2k)^{(2)} d_{\nu(k-1)} + (k-\nu+1)^{(2)} d_{(\nu-1)k},$$
$$\nu = 1, 2, \ldots, k-1, \quad d_{0k} = d_{0(k-1)}.$$

The $d$'s can be calculated from this formula using the values

$$d_{0j} = 1, \quad d_{(j-1)j} = 0, \quad j > 1. \tag{47}$$

To prove these we note that the coefficient of $x^{2k}$ in $B_{2k}(x)$ is 1 and that if $k > 1$, $B_{2k}(x)$ does not contain a term of the first degree in $x$. This is true since

$$B'_{2k}(0) = 2k B_{2k-1}(0) = 0, \quad k > 1.$$

A short table is given to show how the calculations can be carried out. It was made by first filling in the numbers given by (47) and then working downward.

$d_{\nu k}$

| $k \backslash \nu$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 1 | 0 | $B_4$ | | | | | | |
| 3 | 1 | $\frac{1}{2}$ | 0 | $-B_6$ | | | | | |
| 4 | 1 | $\frac{4}{3}$ | $\frac{2}{3}$ | 0 | $B_8$ | | | | |
| 5 | 1 | $\frac{5}{2}$ | 3 | $\frac{3}{2}$ | 0 | $-B_{10}$ | | | |
| 6 | 1 | 4 | $\frac{17}{2}$ | 10 | 5 | 0 | $B_{12}$ | | |
| 7 | 1 | $\frac{35}{6}$ | $\frac{287}{15}$ | $\frac{118}{3}$ | $\frac{691}{15}$ | $\frac{691}{30}$ | 0 | $-B_{14}$ | |
| 8 | 1 | 8 | $\frac{112}{3}$ | $\frac{352}{3}$ | $\frac{718}{3}$ | 280 | 140 | 0 | $B_{16}$ |
| 9 | 1 | $\frac{21}{2}$ | 66 | 293 | $\frac{4557}{5}$ | $\frac{3711}{2}$ | $\frac{10851}{5}$ | $\frac{10851}{10}$ | 0 |

We note from the method of constructing the table that

$$d_{\nu k} > 0, \qquad \nu \leqslant k-2.$$

It is interesting to note that $d_{kk} = (-1)^k B_{2k}$. This results from (45) if we let $x = 0$. Again letting $\phi_{2k}(x) = B_{2k}(x) - B_{2k}$, and noting that $\phi_{2k}(x)$ has no first degree terms we can write

$$\phi_{2k}(x) = (-1)^k \sum_{\nu=0}^{k-2} d_{\nu k}\, z^{k-\nu}, \quad k > 1. \tag{48}$$

By differentiating (48)

$$(-1)^k \phi'_{2k}(x) = (-1)^k 2k\phi_{2k-1}(x) = (1-2x)\sum_{\nu=0}^{k-2} d_{\nu k}(k-\nu)z^{k-\nu-1}$$

$$= (1-2x)(x-x^2)\sum_{\nu=0}^{k-2} d_{\nu k}(k-\nu)z^{k-\nu-2}, \quad k > 1.$$

From this we determine that the maximum value of $(-1)^k\phi_{2k}(x)$ when $0 < x < 1$ is when $x \doteq \frac{1}{2}$. The value of this maximum is given by

$$(-1)^k \phi_{2k}(\tfrac{1}{2}) = \sum_{\nu=0}^{k-2} \frac{d_{\nu k}}{4^{k-\nu}}, \quad k > 1. \tag{49}$$

If we note from (36) that

$$B_{2k}(\tfrac{1}{2}) = \phi_{2k}(\tfrac{1}{2}) + B_{2k} = -\left(1 - \frac{1}{2^{2k-1}}\right) B_{2k},$$

we get $$B_{2k} = (-1)^{k-1}\frac{1}{2-1/2^{2k-1}}\sum_{\nu=0}^{k-2}\frac{d_{\nu k}}{4^{k-\nu}}, \quad k > 1.$$

This is a formula which can be used for the calculation of $B_{2k}$ if the $d_{\nu k}$ are known.

Also simple relations connecting the $d$'s and $B$'s are obtained by formally taking successive derivatives of (45), utilizing (9), and then letting $x = 0$. For example the second derivative yields

$$d_{(k-2)k} = (-1)^k\frac{2k(2k-1)}{2}B_{2k-2}. \tag{50}$$

In general if we write

$$\begin{aligned}(-1)^k(2k)^{(j)}B_{2k-j}(x) &\\ =\sum_{\nu=0}^{k-1} d_{\nu k}[M_{0j}(k-\nu)^{(j)}(1-2x)^j z^{k-\nu-j}-&\\ -M_{1j}(k-\nu)^{(j-1)}(1-2x)^{j-2}z^{k-\nu-j+1}+\ldots+&\\ +(-1)^m M_{mj}(k-\nu)^{(m)}(1-2x)^l z^{k-\nu-j+m}],&\\ j = 2m+l, \qquad l = 0, 1,&\end{aligned} \tag{51}$$

differentiation yields

$$\begin{aligned}&(-1)^k(2k)^{(j+1)}B_{2k-j-1}(x)\\ &=\sum_{\nu=0}^{k-1} d_{\nu k}[M_{0j}(k-\nu)^{(j+1)}(1-2x)^{j+1}z^{k-\nu-j-1}-\\ &-(k-\nu)^{(j)}(1-2x)^{j-1}(M_{1j}+2jM_{0j})z^{k-\nu-j}+\\ &+(k-\nu)^{(j-1)}(1-2x)^{j-3}\{M_{2j}+2(j-2)M_{1j}\}z^{k-\nu+1}-\ldots+\\ &+(-1)^{m+1}2l(k-\nu)^{(m)}M_{mj}\,z^{k-\nu-j+m}].\end{aligned}$$

Expand the left-hand member of this by (51), equate coefficients, and we have

$$\begin{aligned}M_{0(j+1)} &= M_{0j} = 1,\\ M_{1(j+1)} &= M_{1j}+2jM_{0j},\\ M_{2(j+1)} &= M_{2j}+2(j-2)M_{1j},\\ M_{3(j+1)} &= M_{3j}+2(j-4)M_{2j},\\ &\cdots\cdots\cdots\\ M_{m(j+1)} &= M_{mj}+2(l+1)M_{(m-1)j},\\ M_{(m+1)(j+1)} &= 2lM_{mj}.\end{aligned}$$

By means of these relations the coefficients in (51) can be readily calculated from the initial values. For example:

$$M_{0j} = 1, \quad j \geqslant 0.$$
$$M_{10} = 0, \quad M_{11} = 0, \quad M_{12} = 2, \quad M_{13} = 6, \quad M_{14} = 12,$$
$$M_{15} = 20, \quad M_{16} = 30, \quad M_{20} = 0, \quad M_{21} = 0, \quad M_{22} = 0,$$
$$M_{23} = 0, \quad M_{24} = 12, \quad M_{25} = 60, \quad M_{26} = 180.$$

If now in (51) we let $x = 0$ we obtain the following formula:

$$(-1)^k (2k)^{(j)} B_{2k-j} = d_{(k-j)k} M_{0j}(j)^{(j)} - d_{(k-j+1)k} M_{1j}(j-1)^{(j-1)} + \\ + \ldots + (-1)^{(j-2)} d_{(k-2)k} M_{(j-2)j}.$$

This is a relation from which the Bernoulli numbers can be calculated given the $d$'s.

A table of $\delta$'s is given for reference.

| $k \backslash \nu$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| 2 | $\frac{1}{2}$ | 0 | | | | | | | |
| 3 | $\frac{1}{2}$ | $\frac{1}{6}$ | 0 | | | | | | |
| 4 | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{6}$ | 0 | | | | | |
| 5 | $\frac{1}{2}$ | 1 | $\frac{9}{10}$ | $\frac{3}{10}$ | 0 | | | | |
| 6 | $\frac{1}{2}$ | $\frac{5}{3}$ | $\frac{17}{6}$ | $\frac{5}{2}$ | $\frac{5}{6}$ | 0 | | | |
| 7 | $\frac{1}{2}$ | $\frac{5}{2}$ | $\frac{41}{6}$ | $\frac{236}{21}$ | $\frac{691}{70}$ | $\frac{691}{210}$ | 0 | | |
| 8 | $\frac{1}{2}$ | $\frac{7}{2}$ | 14 | $\frac{110}{3}$ | $\frac{359}{6}$ | $\frac{105}{2}$ | $\frac{35}{2}$ | 0 | |
| 9 | $\frac{1}{2}$ | $\frac{14}{3}$ | $\frac{77}{3}$ | $\frac{293}{3}$ | $\frac{1519}{6}$ | $\frac{1237}{3}$ | $\frac{3617}{10}$ | $\frac{3617}{30}$ | 0 |

## 11. The Euler polynomials and numbers

A set of polynomials of much historical importance bears the name of Euler. They can be readily defined in terms of the Bernoulli polynomials as follows. If $E_{n-1}(x)$ denotes the Euler polynomial of degree $n-1$, then

$$E_{n-1}(x) = \frac{2^n}{n}\left[B_n\left(\frac{x+1}{2}\right) - B_n\left(\frac{x}{2}\right)\right], \tag{52}$$

or the equivalent

$$E_{n-1}(x) = \frac{2}{n}\left[B_n(x) - 2^n B_n\left(\frac{x}{2}\right)\right]. \tag{52'}$$

This equivalence is easily proved by means of (34). From this definition it can be proved that

$$\nabla E_n(x) = x^n \tag{53}$$

and it is possible to use (53) as a starting-point for the definition of $E_n(x)$.

From known relations for the Bernoulli polynomials we can prove

$$\frac{d}{dx}E_n(x) = nE_{n-1}(x)$$

and, consequently,

$$\int_x^y E_n(t)\,dt = \frac{E_{n+1}(y) - E_{n+1}(x)}{n+1}.$$

By Taylor's formula,

$$E_n(x+h) = \sum_{\nu=0}^{n} {}_nC_\nu\, h^\nu E_{n-\nu}(x).$$

If in this we let $h = 1$ and use (53), we get the following relation,

$$\sum_{\nu=0}^{n} {}_nC_\nu\, E_\nu(x) + E_n(x) = 2x^n.$$

This can be used to calculate successively the Euler polynomials. We find:

$$E_0(x) = 1, \quad E_1(x) = x - \tfrac{1}{2}, \quad E_2(x) = x(x-1),$$
$$E_3(x) = (x-\tfrac{1}{2})(x^2 - x - \tfrac{1}{2}), \quad E_4(x) = x(x-1)(x^2-x-1),$$
$$E_5(x) = (x-\tfrac{1}{2})(x^4 - 2x^3 - x^2 + 2x + 1),$$
$$E_6(x) = x(x-1)(x^4 - 2x^3 - 2x^2 + 3x + 3),$$
$$\cdots\cdots\cdots\cdots\cdots$$

These polynomials can, of course, also be calculated from the corresponding Bernoulli polynomials by means of (52).

The following relations are readily proved for the Euler polynomials by means of the corresponding relations for the Bernoulli polynomials

$$E_n(1-x) = (-1)^n E_n(x),$$

$$E_n(mx) = m^n \sum_{\nu=0}^{m-1} (-1)^\nu E_\nu\left(x+\frac{\nu}{m}\right), \quad m \text{ odd},$$

$$E_n(mx) = -\frac{2m^n}{n+1} \sum_{\nu=0}^{m-1} (-1)^\nu B_{\nu+1}\left(x+\frac{\nu}{m}\right), \quad m \text{ even}.$$

The Euler polynomials can be made the means of defining a set of numbers known as the Euler numbers, $E_n$. Let

$$E_n(x) = \sum_{\nu=0}^{n} {}_nC_\nu \frac{E_\nu^{(n)}}{2^\nu}(x-\tfrac{1}{2})^{n-\nu}. \tag{54}$$

THEOREM. *The numbers $E_\nu^{(n)}$ are independent of $n$.*

To prove this differentiate (54):

$$\frac{d}{dx}E_n(x) = n \sum_{\nu=0}^{n-1} {}_{n-1}C_\nu \frac{E_\nu^{(n)}}{2^\nu}(x-\tfrac{1}{2})^{n-\nu-1}.$$

But we have already seen that

$$\frac{d}{dx}E_n(x) = nE_{n-1}(x).$$

Expand the right-hand member here by (54). Comparison of coefficients gives us the desired result.

From (52) we readily verify that

$$E_n(0)+E_n(1) = 0.$$

This, together with (54), gives the symbolic relation

$$(E+1)^n+(E-1)^n \doteq 0.$$

If we combine this with $E_0 = 1$ we have a formula from which successive $E_n$'s can be calculated. It also follows that all $E_n$ are integers and that all $E_{2k-1} = 0$.

The first few Euler numbers are as follows:

$$E_0 = 1, \qquad E_2 = -1, \qquad E_4 = 5, \qquad E_6 = -16,$$
$$E_8 = 1385, \qquad E_{10} = -50521.$$

## 12. The Euler polynomials in the interval (0, 1)

We see from (52) that

$$E_{2k}(0) = E_{2k}(1) = E_{2k}(\tfrac{1}{2}) = 0.$$

Also $$E_{2k-1}(0) = -E_{2k-1}(1) = -\frac{2^{2k-2}}{2k-1}B_{2k}(\tfrac{1}{2}) \neq 0.$$

Moreover, the Euler polynomials have no zeros on the interval (0, 1) other than those just discussed. This is readily proved by means of Rolle's theorem, as in the corresponding case for Bernoulli polynomials.

## 13. The Bernoulli and Euler polynomials with difference interval $h$ and the Bernoulli and Euler polynomials of higher order

In our derivation and discussion of the Bernoulli polynomials and numbers we have assumed the difference interval equal to 1. The assumption of a difference interval $h$ is an immediate generalization. The treatment has been given in this book for $h = 1$ inasmuch as this yields the classical and useful theory. However, this restriction is not necessary. The difference interval then appears as an argument of our functions. We proceed exactly as in § 1. Write

$$\phi_n(x, h) = h\sum_{x=0}^{x-h} nx^{n-1}, \quad n > 0,$$

$$\phi_0(x, h) \equiv 0.$$

As formerly we arrive at Bernoulli polynomials and numbers. We now write $B_n(x, h)$ for the Bernoulli polynomials and $B_n(h)$ for the Bernoulli numbers. Certain modifications of our work are, of course, necessary. They are, however, not deemed worthy of the necessary space particularly in view of the very general discussion given in the next section of this chapter and of the discussion in the final sections of Chapter IV and Chapter V. The interested student can refer to Nörlund† and should be able to develop analogous theory to that of this text for himself.

To define the Bernoulli polynomials we apply the operator $h\sum_{x=0}^{x-h}$ to $nx^{n-1}$. There is certainly no reason why this operation

† N. E. Nörlund: *Differenzenrechnung.*

should not be repeated and each time with a different $h$. We let

$$\phi_n(x, h_1, h_2, \ldots, h_e) = h_1 h_2 \ldots h_e \sum_{x_1=0}^{x-h_1} \sum_{x_2=0}^{x_1-h_2} \cdots \sum_{x_e=0}^{x_{e-1}-h_e} n x_e^{n-1}.$$

Proceed as formerly and we arrive at polynomials and numbers which we write

$$B_n(x, h_1, \ldots, h_e) \quad \text{and} \quad B_n(h_1, \ldots, h_e).$$

These are frequently called Bernoulli polynomials and numbers of higher order. A detailed discussion is again not given particularly in view of the next section of this chapter and of the last sections in Chapters IV and V.

The remarks that have just been made with reference to the Bernoulli polynomials and numbers apply equally well to Euler polynomials and numbers. We write

$$E_n(x, h_1, \ldots, h_e) \quad \text{and} \quad E_n(h_1, \ldots, h_e).$$

## 14. Generalizations of the Bernoulli polynomials and numbers

Bernoulli and Euler polynomials are special cases of what is known as Appell polynomial.

*A set of polynomials $A_n(x)$ are called Appell polynomials if they obey the relationship*

$$\frac{d}{dx} A_n(x) = n A_{n-1}(x).$$

Such polynomials have been extensively studied. We propose here, however, a study of sets of polynomials which are far more general than Appell polynomials.

Let us be given two linear operators $P$ and $Q$ with their inverses $P^{-1}$ and $Q^{-1}$. We shall assume that $P$ reduces the degree of any polynomial by 1 and that $Q$ reduces the degree of any polynomial by $k \geqslant 0$, that $P$ operating on a constant gives zero and that $Q$ operating on any polynomial of lesser degree than $k$ gives zero. We assume that $P$, $P^{-1}$, $Q$, $Q^{-1}$ each, where applicable, gives a unique result, except that it is permitted that the result of operating with $P^{-1}$ lack in uniqueness by an arbitrary additive constant and by $Q^{-1}$ by an additive arbitrary polynomial of degree less than $k$.

We assume, moreover, that we are given a set of polynomials $f_n(x)$, where $f_n(x)$ is of degree $n$, such that

$$f_0(x) = 1, \tag{55}$$

$$Pf_n(x) = nf_{n-1}(x), \tag{56}$$

$$\begin{aligned} &Q^{-1}n(n-1)\dots(n-k+1)f_{n-k}(x) \\ &\quad = {}_nL_0 f_n(x)+n\,{}_nL_1 f_{n-1}(x)+{}_nC_2\,{}_nL_2 f_{n-2}(x)+ \\ &\qquad + {}_nC_k\,{}_nL_{n-k} f_k(x)+c_1 x^{k-1}+c_2 x^{k-2}+\dots+c_k, \quad 0 \leqslant k \leqslant n, \end{aligned} \tag{57}$$

where ${}_nL_0,\dots, {}_nL_{n-k}$ are independent of $x$ but uniquely determined. The $c$'s may be determined constants or may be arbitrary depending upon the nature of the operator $Q^{-1}$.

From (57)

$$\begin{aligned} &\frac{1}{n}PQ^{-1}n(n-1)\dots(n-k+1)f_{n-k}(x) \\ &\quad = {}_nL_0 f_{n-1}(x)+(n-1)\,{}_nL_1 f_{n-2}(x)+\frac{(n-1)(n-2)}{2!}\,{}_nL_2 f_{n-3}(x)+ \\ &\qquad +\dots+{}_{n-1}C_k\,{}_nL_{n-k} f_{k-1}(x)+(\text{terms of degree less than } k-1). \end{aligned} \tag{58}$$

However,

$$\begin{aligned} &\frac{1}{n}Q^{-1}Pn(n-1)\dots(n-k+1)f_{n-k}(x) \\ &\quad = Q^{-1}(n-1)(n-2)\dots(n-k)f_{n-k-1}(x) \\ &\quad = {}_{n-1}L_0 f_{n-1}(x)+(n-1)\,{}_{n-1}L_1 f_{n-2}(x)+{}_nC_2\,{}_{n-1}L_2 f_{n-3}(x)+ \\ &\qquad +\dots+{}_{n-1}C_k\,{}_{n-1}L_{n-k} f_{k-1}(x)+(\text{terms of degree less than } k). \end{aligned} \tag{59}$$

We next assume

$$PQ^{-1}f_{n-k}(x) = Q^{-1}Pf_{n-k}(x)+(\text{terms of degree less than } k).$$

Under this assumption

$$\begin{aligned} &{}_nL_0 f_{n-1}(x)+(n-1)\,{}_nL_1 f_{n-2}(x)+{}_{n-1}C_2\,{}_nL_2 f_{n-3}(x)+ \\ &\qquad\qquad +\dots+{}_{n-1}C_k\,{}_nL_{n-k} f_{k-1}(x) \\ &\quad \equiv {}_{n-1}L_0 f_{n-1}(x)+(n-1)\,{}_{n-1}L_1 f_{n-2}(x)+{}_{n-1}C_2\,{}_{n-1}L_2 f_{n-3}(x)+ \\ &\qquad +\dots+{}_{n-1}C_k\,{}_{n-1}L_{n-k} f_{n-2}(x)+(\text{terms of degree less than } k). \end{aligned} \tag{60}$$

Equating coefficients we find

$$ {}_nL_0 = {}_{n-1}L_0, \quad {}_nL_1 = {}_{n-1}L_1, \quad {}_nL_2 = {}_{n-1}L_2, \quad \dots . $$

In other words, the $L$'s are independent of $n$. Henceforth we write simply $L_0$, $L_1$, $L_2$,... . We let

$$ F_n(x) = L_0 f_n(x) + nL_1 f_{n-1}(x) + {}_nC_2 L_2 f_{n-2}(x) + \\ + \dots + L_n f_0(x), \quad n \geqslant 0. \quad (61) $$

By (57)

$$ QF_n(x) = n(n-1)\dots(n-k+1)f_{n-k}(x), \quad n \geqslant k. \quad (62) $$

For reasons of symmetry and consequent simplification of the sequel we do not take the $L_n$'s just defined as our fundamental sequence of numbers but the set $g_n$ determined from the equations $f_n(g) = L_n$, where subscripts are applied to $g$ rather than exponents in the expansion of $f_n(g)$. Such a determination is always possible and unique. We then write

$$ F_n(x) = f_0(g)f_n(x) + nf_1(g)f_{n-1}(x) + \\ + {}_nC_2 f_2(g) f_{n-2}(x) + \dots + f_n(g) f_0(x), \quad n \geqslant 0. \quad (63) $$

We choose to write this $f_n(x+g)$ of which more will be said later. We then have

$$ F_n(x) = f_n(x+g), \quad (64) $$

$$ PF_n(x) = nF_{n-1}(x). \quad (65) $$

The polynomials $F(x)$ are the polynomials† in which we are interested and the numbers $g_n$ constitute the corresponding sequence of numbers.

*Special cases:* We now consider these special cases.

(*a*) Bernoulli polynomials and numbers:

$$ P = \frac{d}{dx}, \qquad Q = \Delta, \qquad f_n(x) = x^n. $$

(*b*) Bernoulli polynomials and numbers of the second kind:‡

$$ P = \underset{h}{\Delta}, \qquad Q = \frac{d}{dx}, \qquad f_n(x) = x^{(n)}. $$

† If we replace (56) by $Pf_n(x) = n(n-1)\dots(n-h+1)f_{n-h}(x)$ and the requirement that the result of operating with $P$ on any polynomial of degree less than $h$ is zero, we are led to sequences of polynomials each of degree differing from that of the previous by $h$, a somewhat more general situation than that treated in the text.

‡ See C. Jordan: *Calculus of Finite Differences*, p. 265.

(*c*) Bernoulli polynomials and numbers of higher order as defined by Nörlund:†

$$P = \frac{d}{dx}, \qquad Q = \underset{h_1 h_2 \dots h_k}{\Delta^k}, \qquad f_n(x) = x^n, \qquad k > 1.$$

(*d*) Bernoulli polynomials and numbers of the second kind of higher order:‡

$$P = \underset{h}{\Delta}, \qquad Q = \frac{d^k}{dx^k}, \qquad f_n(x) = x^{(n)}, \qquad k > 1.$$

(*e*) Bernoulli polynomials with Bernoulli numbers of higher order as defined by Vandiver:§

$$P = \frac{d}{dx}, \qquad Q = \underset{h_1 h_2 \dots h_k}{\Delta^k}, \qquad f_n(x) = (x+l)^n.$$

Vandiver's numbers are the $L_n$'s of the text rather than the $g_n$'s.

(*f*) Euler polynomials:

$$P = \frac{d}{dx}, \qquad Q = \nabla, \qquad f(x) = x^n.$$

(*g*) Euler polynomials of higher order:

$$P = \frac{d}{dx}, \qquad Q = \nabla^k, \qquad f(x) = x^n.$$

(*h*) $$P = \frac{d}{dx}, \qquad Q = \frac{\Delta^k}{h_1 h_2 \dots h_k},$$

$f_n(x)$ any set of Appell‖ polynomials.

(*i*) $$P = \Delta, \qquad Q = \frac{d^k}{dx^k},$$

$f(x)$ any set of Appell polynomials of the second kind.

(*j*) $$P = a_0 D^r + a_1 D^{r-1} + \dots + a_{r-1} D = R(D),$$

where $D = d/dx$ and $a_0, a_1, \dots, a_{r-1}$ are constants: $a_{r-1} \neq 0$.

$$Q = b_0 \Delta^s + b_1 \Delta^{s-1} + \dots + b_s, \quad b_0, b_1, \dots, b_s \text{ being constants.}$$

† N. E. Nörlund: *Differenzenrechnung*, p. 119.

‡ This name is applied through analogy with the Bernoulli polynomials of higher order as defined by Nörlund and as a natural extension of the Bernoulli polynomials of the second kind already referred to.

§ H. S. Vandiver: *Proceedings of the National Academy of Science*, **23**, 555.

‖ We understand by a set of Appell polynomials a set of polynomials such that $(d/dx)f_n(x) = nf_{n-1}(x)$ and of the second kind such that $\Delta\phi_n(x) = hn\phi_{n-1}(x)$.

$f_n(x)$ are polynomials obtained by successive solutions of the equation

$$R(D)f_n(x) = nf_{n-1}(x).$$

(*k*) Interchange $D$ and $\Delta$ in (*j*).

*Table of Bernoulli Numbers*

$B_1 = -1/2$
$B_2 = 1/6$
$B_4 = -1/30$
$B_6 = 1/42$
$B_8 = -1/30$
$B_{10} = 5/66$
$B_{12} = -691/2730$
$B_{14} = 7/6$
$B_{16} = -3617/510$
$B_{18} = 43867/798$
$B_{20} = -174611/330$
$B_{22} = 854513/138$
$B_{24} = -236364091/2730$
$B_{26} = 8553103/6$
$B_{28} = -23749461029/870$
$B_{30} = 8615841276005/14322$
$B_{32} = -7709321041217/510$
$B_{34} = 2577687858367/6$
$B_{36} = -26315271553053477373/1919190$
$B_{38} = 2929993913841559/6$
$B_{40} = -261082718496449122051/13530$
$B_{42} = 1520097643918070802691/1806$
$B_{44} = -27833269579301024235023/690$
$B_{46} = 596451111593912163277961/282$
$B_{48} = -5609403368997817686249127547/46410$
$B_{50} = 495057205241079648212477525/66$
$B_{52} = -801165718135489957347924991853/1590$
$B_{54} = 29149963634884862421418123812691/798$
$B_{56} = -2479392929313226753685415739663229/870$
$B_{58} = 84483613348880041862046775994036021/354$
$B_{60} = -1215233140483755572040304994079820246041491/56786730$

## EXERCISES

1. By elementary calculus prove

$$|B_{2k}(x)| < |B_{2k}|, \quad 0 < x < 1.$$

2. Prove $$\nabla^{-1}B_n(x) = B_n(x) - \frac{n}{2}E_{n-1}(x).$$

3. Prove

$$\frac{24(2n+1)(2n+2)}{(2\pi)^4} < \left|\frac{B_{2n+2}}{B_{2n}}\right| < \frac{(2n+1)(2n+2)}{(2\pi)^2}.$$

4. Prove

$$(-1)^n\frac{B_n}{n!} = \begin{vmatrix} \frac{1}{2!} & 1 & 0 & 0 & \cdots & 0 \\ \frac{1}{3!} & \frac{1}{2!} & 1 & 0 & \cdots & 0 \\ \frac{1}{4!} & \frac{1}{3!} & \frac{1}{2!} & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ \frac{1}{(n+1)!} & \frac{1}{n!} & \frac{1}{(n-1)!} & & \cdots & \frac{1}{2!} \end{vmatrix}$$

5. Prove $B_n(x)+\frac{1}{2^n}x^{n-1}$ an even function when $n$ is even and an odd function when $n$ is odd.

6. Express the Bernoulli polynomials as sums of factorials.

7. If Bernoulli polynomials of the second kind are defined as are Bernoulli polynomials only interchanging $\Delta$ and $D$, $\sum$ and $\int$, $x^{(n)}$ and $x^n$, develop a theory for Bernoulli polynomials and numbers of the second kind analogous to that developed in the text for Bernoulli polynomials and numbers.

8. Express the Bernoulli polynomials of the second kind as a sum of factorials.

9. Calculate the first ten Bernoulli polynomials.

10. Calculate the first fifteen Bernoulli numbers.

11. Obtain an approximate value for $B_{70}$. Discuss your error.

12. Calculate the first five Bernoulli polynomials of the second kind.

13. Work out

$$\sum(x^9+11x^8+6x^7-4x^5), \quad h=1.$$

Use Bernoulli polynomials.

14. Work out $\int(x^{(9)}+11x^{(8)}+6x^{(7)}-4x^{(5)})\,dx.$

Use Bernoulli polynomials of the second kind.

15. Express Bernoulli polynomials in terms of Bernoulli polynomials of the second kind.

16. Express Bernoulli polynomials of the second kind in terms of Bernoulli polynomials.

17. Express the Euler polynomials as sums of factorials.

18. Obtain a Fourier expansion for the Euler polynomials.

# IV

# SUMMATION FORMULAE

## 1. The Euler-Maclaurin summation formulae

In the previous chapter we obtained formula (29) which we denoted by the name Euler-Maclaurin summation formula. As remarked there, this formula when used for a polynomial of degree $m$ is an identity. However, the Euler-Maclaurin formula, as given by (29) of Chapter III, permits of generalization, as will be immediately illustrated.

Let $F(x)$ be a function for which the first $m$ derivatives exist and are continuous at all points of the interval under consideration. Let $\bar{B}_p(x)$ represent the function of period 1 which coincides with $B_p(x)$ over the closed interval $0 \leqslant x \leqslant 1$, $p = 1, 2,\dots$. These functions, except $\bar{B}_1(x)$, are continuous for all real values of $x$. This latter function has a finite discontinuity at

$$x = 0, 1, 2,\dots.$$

We now let the difference interval be $h$. Let $\omega$ be a parameter, $0 \leqslant \omega \leqslant 1$, and then let

$$-R_m(x) = h^m \int_0^1 \frac{\bar{B}_m(\omega-t)}{m!} F^{(m)}(x+ht)\, dt. \tag{1}$$

With this formula in view we rewrite (9) of the last chapter

$$\frac{d}{dx}\bar{B}_n(x) = n\bar{B}_{n-1}(x), \quad 0 < x < 1, \tag{2}$$

and apply integration by parts repeatedly to (1). We get

$$-R_m(x) = \sum_{\nu=2}^{m} \frac{h^{\nu-1}}{\nu!} B_\nu(\omega)\Delta F^{(\nu-1)}(x) + h\int_0^1 \bar{B}_1(\omega-t)F'(x+ht)\, dt. \tag{3}$$

Furthermore

$$\int_0^1 \bar{B}_1(\omega-t)F'(x+ht)\, dt = \int_{\omega-1}^{\omega} \bar{B}_1(y)F'\{x+h(\omega-y)\}\, dy$$

$$= \int_{\omega-1}^{0} (y+\tfrac{1}{2})F'\{x+h(\omega-y)\}\, dy + \int^{\omega} (y-\tfrac{1}{2})F'\{x+h(\omega-y)\}\, dy$$

$$= -\frac{1}{h}F(x+\omega h)+\frac{(\omega-\frac{1}{2})}{h}\Delta F(x)+\frac{1}{h}\int\limits_{\omega-1}^{\omega} F\{x+h(\omega-y)\}\,dy$$

$$= -\frac{1}{h}F(x+\omega h)+\frac{B_1(\omega)}{h}\Delta F(x)+\frac{1}{h}\int\limits_{\omega-1}^{\omega} F\{x+h(\omega-y)\}\,dy.$$

Substitute this in (3). Replace $-R_m(x)$ by the right-hand member of (1) and transpose. We have, after changing the variable in the first integral,

$$F(x+\omega h) = \frac{1}{h}\int\limits_{x}^{x+h} F(t)\,dt + \sum_{\nu=1}^{m}\frac{h^{\nu-1}}{\nu!}B_\nu(\omega)\Delta F^{(\nu-1)}(x) - $$
$$-h^m\int\limits_0^1 \frac{\bar{B}_m(\omega-t)}{m!}F^{(m)}(x+ht)\,dt. \quad (4)$$

This is the Euler-Maclaurin formula sought.

## 2. The Euler summation formula

If we write (4) for $x = a, a+h, a+2h,\ldots, a+(n-1)h, b-h$ and sum, we get

$$\sum_{x=a}^{b-h} F(x+\omega h)$$
$$= \frac{1}{h}\int\limits_a^b F(t)\,dt + \sum_{\nu=1}^{m}\frac{h^{\nu-1}}{\nu!}B_\nu(\omega)[F^{(\nu-1)}(b)-F^{(\nu-1)}(a)] -$$
$$-h^m\int\limits_0^1 \frac{\bar{B}_m(\omega-t)}{m!}\left[\sum_{x=a}^{b-h} F^{(m)}(x+ht)\right]dt. \quad (5)$$

This formula is sometimes called the Euler-Maclaurin summation formula instead of (4). It is, however, more usually known simply as the *Euler summation formula* and will be so designated by us.

## 3. The Euler formula when $\omega = 0$

In (5) let $\omega = 0$. Let $\bar{\phi}_m(t) \equiv \bar{B}_m(t)-B_m$ and replace $\bar{B}_m(-t)$ by $\bar{\phi}_m(-t)+B_m \equiv \bar{\phi}_m(1-t)+B_m$. Let $m$ be even, $m = 2k$, and

note that $\bar{\phi}_{2k}(1-t) = \bar{\phi}_{2k}(t)$ and that $B_{2k-1} = 0$, $k > 1$. Formula (5) then reduces to

$$\sum_{x=a}^{b-h} F(x) = \frac{1}{h}\int_a^b F(t)\,dt + \sum_{\nu=1}^{2k-2} \frac{h^{\nu-1}}{\nu!} B_\nu[F^{(\nu-1)}(b) - F^{(\nu-1)}(a)] -$$
$$-\frac{h^{2k}}{(2k)!}\int_0^1 \phi_{2k}(t)\Big[\sum_{x=a}^{b-h} F^{(2k)}(x+ht)\Big]\,dt, \quad k > 1. \tag{6}$$

As we know, $\phi_{2k}(t)$ retains the same sign over the interval (0, 1) and hence we can apply the first Theorem of the Mean for integrals in (6). Denote the last expression on the right by $R'_{2k}(a, b)$. We get

$$R'_{2k}(a,b) = -\frac{h^{2k}}{(2k)!}\int_0^1 \phi_{2k}(t)\Big[\sum_{x=a}^{b-h} F^{(2k)}(x+ht)\Big]\,dt$$
$$= -\frac{h^{2k}}{(2k)!}\sum_{x=a}^{b-h} F^{(2k)}(x+\theta h)\int_0^1 \phi_{2k}(t)\,dt, \quad 0 < \theta < 1. \tag{7}$$

But
$$\phi_{2k}(t) = \frac{1}{2k+1}\phi'_{2k+1}(t) - B_{2k}.$$

Hence, since
$$\phi_{2k+1}(1) = \phi_{2k+1}(0) = 0,$$
$$\int_0^1 \phi_{2k}(t)\,dt = -B_{2k}.$$

Consequently,
$$R'_{2k}(a,b) = \frac{h^{2k}}{(2k)!}B_{2k}\sum_{x=a}^{b-h} F^{(2k)}(x+\theta h). \tag{8}$$

This is a convenient form for the remainder when $\omega = 0$ and $m = 2k$. It is also the most important case.

## 4. The Euler-Maclaurin formula again

From (6) and (8) we obtain the Euler-Maclaurin summation formula in the following form:

$$F(x) = \frac{1}{h}\int_x^{x+h} F(t)\,dt + \sum_{\nu=1}^{2k-2}\frac{h^{\nu-1}}{\nu!}B_\nu \Delta F^{(\nu-1)}(x) +$$
$$+\frac{h^{2k}}{(2k)!}B_{2k}F^{(2k)}(x+\theta h), \quad 0 < \theta < 1. \tag{9}$$

It is to be noted that the Euler-Maclaurin formula appears as a special case of the Euler summation formula.

## 5. Another form of the remainder in the Euler formula

Assume that $F_2^{(j)}(t) > 0$, $j \geqslant K$, each retains the same sign for $a \leqslant t \leqslant b$ and that $F^{(2k)}(t)F^{(2k-2)}(t) > 0$. Consider the middle member in (7). Integrating by parts twice, we obtain

$$R'_{2k}(a,b) = \frac{-h^{2k}}{(2k)!}\int\limits_0^1 \phi_{2k}(t)\sum_{x=a}^{b-h} F^{(2k)}(x+ht)\,dt$$

$$= -\frac{h^{2k-3}}{(2k-2)!}B_{2k-2}[F^{(2k-3)}(b)-F^{(2k-3)}(a)]-$$

$$-\frac{h^{2k-2}}{(2k-2)!}\int\limits_0^1 \phi_{2k-2}(t)\sum_{x=a}^{b-h} F^{(2k-2)}(x+ht)\,dt. \tag{10}$$

Now if $A = B+C$ and $CA < 0$ then

$$A = \theta B, \quad 0 < \theta < 1. \tag{11}$$

We wish to apply this to (10). Note that $\phi_{2k}(t)\phi_{2k-2}(t) < 0$. We have

$$R'_{2k}(a,b) = -\theta\frac{h^{2k-3}}{(2k-2)!}B_{2k-2}[F^{(2k-3)}(b)-F^{(2k-3)}(a)]. \tag{12}$$

Similarly, assuming the existence and continuity of $F^{(2k+2)}(t)$, if integration by parts is applied twice to

$$-\frac{h^{2k+2}}{(2k+2)!}\int\limits_0^1 \phi_{2k+2}(t)\sum_{x=a}^{b-h} F^{(2k+2)}(x+ht)\,dt, \tag{13}$$

we obtain

$$R'_{2k}(a,b) = \theta\frac{h^{2k-1}}{(2k)!}B_{2k}[F^{(2k-1)}(b)-F^{(2k-1)}(a)]. \tag{14}$$

## 6. Another form for the Euler formula

We assume that $F(t)$ obeys the following conditions which we call *conditions* A.

(*a*) $F^{(2k)}(t)$ retains the same sign for all positive values of $t$.

(b) $F^{(2k)}(t).F^{(2k-2)}(t) > 0,\ t > t_0$.

(c) $F^{(m)}(t) \to 0$ when $t \to \infty,\ m > K$.

We also assume $\omega = 0$.

Consider the remainder $R'_{2k}(a, b)$, as given by the middle member in (7). Since $\phi_{2k}(t)$ retains the same sign $0 < t < 1$, this increases in absolute value, retaining a fixed sign, as $b$ increases. But, by (14),

$$|R'_{2k}(a,b)| < \frac{h^{2k-1}}{(2k)!}[|F^{(2k-1)}(b)|+|F^{(2k-1)}(a)|]|B_{2k}|.$$

If $b$ is sufficiently large, $b > B$, since $F^{(2k-1)}(b) \to 0$ when $b \to \infty$,

$$|R'_{2k}(a,b)| < \frac{h^{2k}}{(2k)!}[1+F^{(2k-1)}(a)]|B_{2k}|,$$

which is independent of $b$. Hence $R'_{2k}(a, b)$ approaches a limit when $b \to \infty$. Denote this limit by $\rho_{2k}(a)$. If $c > b$ we have by (8)

$$R'_{2k}(a,b) = R'_{2k}(a,c)-R'_{2k}(b,c).$$

Let $c \to \infty$ and we have

$$R'_{2k}(a,b) = \rho_{2k}(a)-\rho_{2k}(b).$$

We now let

$$C_k = \rho_{2k}(a)-\sum_{\nu=1}^{2k-2}\frac{h^{\nu-1}}{\nu!}B_\nu F^{(\nu-1)}(a). \tag{15}$$

We then have, by (6),

$$\sum_{x=a}^{b-h} F(x) = \frac{1}{h}\int_a^b F(t)\,dt+C_k+\sum_{\nu=1}^{2k-2}\frac{h^{\nu-1}}{\nu!}B_\nu F^{(\nu-1)}(b)-\rho_{2k}(b). \tag{16}$$

Now let $b \to \infty$ in (14) and then replace $a$ by $b$. We have

$$\rho_{2k}(b) = -\theta\frac{h^{2k-1}}{(2k)!}B_{2k}F^{(2k-1)}(b), \quad 0 < \theta < 1. \tag{17}$$

We now readily prove that $C_k$ is independent of $k$ when $2k-1 > K$. To do this equate right-hand members in (16) for two different values of $k$ such that in both cases $2k-1 > K$, and allow $b$ to become infinite. The result is immediate. Denote

$C_k$ by $C$. We then have arrived at the following formula under conditions A on $F(x)$:

$$\sum_{x=a}^{b-h} F(x) = \frac{1}{h}\int_a^b F(t)\,dt + C + \sum_{\nu=1}^{2k-2} \frac{h^{\nu-1}}{\nu!} B_\nu F^{(\nu-1)}(b) +$$

$$+\theta\frac{h^{2k-1}}{(2k)!} B_{2k} F^{(2k-1)}(b), \quad 0 < \theta < 1, \quad 2k-1 > K > 1. \quad (18)$$

This proves to be a most useful form of the Euler formula.

If we use (12) instead of (14) in the derivation of (18) we get

$$\rho_{2k}(b) = \theta\frac{h^{2k-3}}{(2k-2)!} B_{2k-2} F^{(2k-3)}(b), \quad 0 < \theta < 1. \quad (19)$$

If in (18) we let $h = 1$, $a = 0$, and $b = 1$ and transpose we get

$$\int_0^1 F(t)\,dt = F(0) - C - \sum_{\nu=1}^{2k-2} \frac{h^{\nu-1}}{\nu!} B_\nu F^{(\nu-1)}(1) +$$

$$+\theta\frac{h^{2k-1}}{(2k)!} B_{2k} F^{(2k-1)}(0), \quad 0 < \theta < 1.$$

More generally from (9) we obtain

$$\int_0^1 F(t)\,dt = F(0) - \sum_{\nu=1}^{2k-1} \frac{h^{\nu-1}}{\nu!} B_\nu \Delta F^{(\nu-1)}(0) -$$

$$-\frac{h^{2k}}{(2k)!} B_{2k} F^{(2k)}(\theta h), \quad 0 < \theta < 1. \quad (20)$$

These are useful formulae for approximating a definite integral.

## 7. A generating function for the Bernoulli polynomials

Let $$F(x) = e^{tx}.$$

Apply formula (4) to $F(x+\omega)$ and let $x = 0$ and $h = 1$. We get

$$e^{t\omega} = \frac{e^t-1}{t} + \sum_{\nu=1}^{m} \frac{B_\nu(\omega)}{\nu!} t^{\nu-1}(e^t-1) - t^m \int_0^1 \frac{\bar{B}_m(\omega-\alpha)}{m!} e^{\alpha t}\,d\alpha.$$

But, by (43) and (44) of the last chapter,

$$\left| t^m \int_0^1 \frac{\bar{B}_m(\omega-\alpha)}{m!} e^{\alpha t}\,d\alpha \right| \leqslant 2\frac{t^{m-1}}{(2\pi)^m}[e^t-1]\frac{1}{1-1/2^{m-1}}.$$

This approaches zero when $m$ becomes infinite if $|t| < 2\pi$. Hence letting $m \to \infty$ and dividing by $(e^t-1)/t$ under the assumption that $t > 0$ we have

$$\frac{te^{t\omega}}{e^t-1} = 1+\sum_{\nu=1}^{\infty} \frac{B_\nu(\omega)}{\nu!} t^\nu, \quad 0 < t < 2\pi. \tag{21}$$

This formula might well be taken as the means of defining the Bernoulli polynomials. It can be obtained by developing $te^{t\omega}/(e^t-1)$ by Maclaurin's series.

If in (21) we let $\omega = 0$, we get a corresponding formula for the Bernoulli numbers, namely,

$$\frac{t}{e^t-1} = 1+\sum_{\nu=1}^{\infty} \frac{B_\nu}{\nu!} t^\nu = e^{Bt}, \quad 0 < t < 2\pi. \tag{22}$$

Here, of course, exponents are applied to $t$ and subscripts to $B$.

## 8. Theorem of von Staudt

Formula (22) can be made the starting-point for a theorem on the Bernoulli numbers known as the Theorem of von Staudt.

Denote by $D_0 f(t)$ the derivative of $f(t)$ with respect to $t$ with $t = 0$. Then by (22), $k > 0$,

$$B_{2k} = D_0^{2k}\left(\frac{t}{e^t-1}\right) = -D_0^{2k}\frac{\log\{1-(1-e^t)\}}{1-e^t}.$$

If $t$ is so small that $|1-e^t| < 1$ we can write

$$B_{2k} = D_0^{2k} \sum_{\lambda=1}^{\infty} \frac{(1-e^t)^{\lambda-1}}{\lambda}.$$

The power series can be differentiated term by term, and

$$D_0^{2k}(1-e^t)^{\lambda-1} = 0 \quad \text{if} \quad \lambda > 2k+1.$$

Consequently, we have

$$B_{2k} = \sum_{\lambda=1}^{2k+1} D_0^{2k}\frac{(1-e^t)^{\lambda-1}}{\lambda}. \tag{23}$$

We proceed now to consider (23) in detail.

First, suppose that $\lambda = ab$ is a composite number greater than or equal to 6, $a \geqslant 2$, $b \geqslant 2$. Then let

$$\lambda - 1 = a+b+c \quad (\text{here } c \geqslant 0).$$

Hence $$(1-e^t)^{\lambda-1} = (1-e^t)^a(1-e^t)^b(1-e^t)^c.$$

Now form $$D_0^{2k}(1-e^t)^a(1-e^t)^b(1-e^t)^c.$$

Perform the differentiation by the product rule. All non-vanishing terms will contain $(a!)(b!)(c!)$ as a factor. Hence

$$D_0^{2k}(1-e^t)^{\lambda-1} \equiv 0 \pmod{ab}.$$

Next, by carrying out the differentiation,

$$D_0^{2k}(1-e^t)^3 \equiv -3.1^{2k}+3.2^{2k}-3^{2k} \equiv 1+0-1 \equiv 0 \pmod 4.$$

Consequently, when $\lambda$ is a composite number greater than or equal to 4 the corresponding term in (23) will give an integer as contribution to the total sum.

Now suppose $\lambda = p$ to be a prime. We then can write

$$2k = q(p-1)+r, \quad 0 \leqslant r < p-1.$$

By employing the binomial formula and carrying out the differentiation we find that

$$D_0^{2k}(1-e^t)^{p-1}$$
$$= -{}_{p-1}C_1\, 1^{2k} + {}_{p-1}C_2\, 2^{2k} - \dots + (-1)^{p-1}{}_{p-1}C_{p-1}(p-1)^{2k}. \quad (24)$$

Moreover

$$D_0^r(1-e^t)^{p-1} = \begin{cases} -{}_{p-1}C_1\, 1^r + {}_{p-1}C_2\, 2^r - \dots + \\ \qquad + (-1)^{p-1}{}_{p-1}C_{p-1}(p-1)^r, & r \neq 0, \\ 1 - {}_{p-1}C_1\, 1^r + {}_{p-1}C_2\, 2^r - \dots + \\ \qquad + (-1)^{p-1}{}_{p-1}C_{p-1}(p-1)^r, & r = 0. \end{cases} \quad (25)$$

Now $$m^{2k} = (m^{p-1})^q m^r.$$

Since by Fermat's first theorem, if $1 \leqslant m < p$,

$$m^{p-1} \equiv 1 \pmod p,$$

we have $$(m^{p-1})^q \equiv 1 \pmod p.$$

Consequently,

$$m^{2k} = (m^{p-1})^q m^r \equiv m^r \pmod p, \quad 1 \leqslant m < p.$$

Hence from (24) and (25)

$$D_0^{2k}(1-e^t)^{p-1} \equiv D_0^r(1-e^t)^{p-1}, \quad D_0^r(1-e^t)^{p-1}-1 \pmod{p} \quad \begin{matrix} r \neq 0, \\ r = 0. \end{matrix}$$

But $D_0^r(1-e^t)^{p-1} = 0$. Hence

$$D_0^{2k}(1-e^t)^{p-1} \equiv \begin{cases} 0 \\ -1 \end{cases} \pmod{p}, \quad \begin{matrix} r \neq 0, \\ r = 0. \end{matrix}$$

Consequently, if $2k$ is not divisible by $p-1$, the contribution of the corresponding term in (25) is an integer. If, however, $2k$ is divisible by $p-1$ the contribution is an integer diminished by $1/p$. We consequently have proved the following relation which bears the name Theorem of von Staudt:

$$B_{2k} = G_{2k} - \sum \frac{1}{p}, \tag{26}$$

where $G_{2k}$ is an integer, and the summation is to be taken for all primes for which $p < 2k$ and $(p-1)$ is a factor of $2k$.

## 9. Power-series developments for $\cot x$ and $\tan x$

Let $F(x) = \cos(tx)$, with $h = 1$, apply (9), and let $x = 0$. We have

$$1 = \frac{\sin t}{t} + B_1(\cos t - 1) - \frac{B_2}{2!} t \sin t + \frac{B_4}{4!} t^3 \sin t - \ldots +$$
$$+(-1)^{k-1} \frac{B_{2k-2}}{(2k-2)!} t^{2k-3} \sin t + (-1)^{k-1} t^{2k-1} \frac{B_{2k}}{(2k)!} \cos(\theta t),$$
$$0 < \theta < 1.$$

Here again, by the use of (43) of the last chapter, one readily shows that

$$t^{2k} \frac{B_{2k}}{(2k)!} \cos(\theta t)$$

approaches zero, as $k \to \infty$ if $|t| < 2\pi$. Consequently, if

$$0 < |t| < 2\pi,$$

$$1 = \frac{\sin t}{t} + B_1(\cos t - 1) - \frac{B_2}{2!} t \sin t + \frac{B_4}{4!} t^3 \sin t - \ldots +$$
$$+(-1)^{k-1} \frac{B_{2k-2}}{(2k-2)!} t^{2k-3} \sin t + \ldots. \tag{27}$$

Replace $1-B_1(\cos t-1)$ by $\cos^2\frac{1}{2}t$ and then divide (27) through by $\sin t = 2\sin\frac{1}{2}t\cos\frac{1}{2}t$ and we get

$$\tfrac{1}{2}\cot\tfrac{1}{2}t = \frac{1}{t}-\frac{B_2}{2!}t+\frac{B_4}{4!}t^3-\frac{B_6}{6!}t^5+\dots.$$

Replace $\frac{1}{2}t$ by $x$ and we have

$$\cot x = \frac{1}{x}-\frac{2^2B_2}{2!}x+\frac{2^4B_4}{4!}x^3-\frac{2^6B_6}{6!}x^5+\dots \tag{28}$$

convergent if $0 < x < \pi$.

With the help of the relation

$$\tan x = \cot x - 2\cot 2x$$

we readily find

$$\tan x = 2^2(2^2-1)\frac{B_2}{2!}x-2^4(2^4-1)\frac{B_4}{4!}x^3+2^6(2^6-1)\frac{B_6}{6!}x^5-\dots$$

convergent when $|x| < \frac{1}{2}\pi$.

### 10. The sum $\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\dots+\frac{1}{x-1}$

Let $F(x) = 1/x$, $a = 1$, $h = 1$, $b = x$, and apply formula (18). We get

$$1+\frac{1}{2}+\frac{1}{3}+\dots+\frac{1}{x-1} = C+\log x-\frac{1}{2x}-\frac{B_2}{2x^2}-\frac{B_4}{4x^4}-\dots-$$
$$-\frac{B_{2k-2}}{(2k-2)x^{2k-2}}-\theta\frac{B_{2k}}{2kx^{2k}},\quad 0<\theta<1. \tag{29}$$

If we allow $x$ to become infinite,

$$C = \lim_{x\to\infty}\left[\sum_{x=1}^{x-1}\frac{1}{x}-\log x\right]$$

appears as Euler's constant, which we denote by $\gamma$. As a matter of fact, formula (29) is a convenient formula for the calculation of $\gamma$. If we take $x = 10$ and $k = 7$ we find

$$\gamma = 0{\cdot}577215664901532\dots.$$

If for a given value of $x$, $k$ is made to become infinite, the resulting infinite series will diverge as is proved by the test-ratio test.

## 11. Stirling's series

Let $F(x) = \log x$, $a = 1$, $h = 1$, $b = x$, and apply formula (18),

$$\sum_{x=1}^{x-1} \log x = \log\{(x-1)!\} = x\log x - x + C + B_1 \log x +$$
$$+\frac{B_2}{2!}x^{-1}+\frac{B_4}{4!}(2!)x^{-3}+\ldots+\frac{B_{2k-2}}{(2k-2)!}\{(2k-4)!\}x^{-(2k-3)}+$$
$$+\theta\frac{B_{2k}}{(2k)!}\{(2k-2)!\}x^{-(2k-1)}, \quad 0<\theta<1. \tag{30}$$

If to both sides of this equation we add $\log x$ and replace $B_1$ by $-\frac{1}{2}$ we have

$$\log(x!) = C+(x+\tfrac{1}{2})\log x - x + \frac{B_2}{1.2}x^{-1}+\frac{B_4}{3.4}x^{-3}+\ldots+$$
$$+\frac{B_{2k-2}}{(2k-3)(2k-2)}x^{-(2k-3)}+\theta\frac{B_{2k}}{(2k-1)(2k)}x^{-(2k-1)},$$
$$0<\theta<1. \tag{31}$$

This formula is valid when $k > 1$. If $k = 2$ the remainder term can be modified by (19) and we write

$$\log(x!) = C+(x+\tfrac{1}{2})\log x - x+\theta\frac{B_2}{2}x^{-1}. \tag{32}$$

It remains to evaluate $C$. To this end, we introduce Wallis's formula for $\frac{1}{2}\pi$:

$$\tfrac{1}{2}\pi = \lim_{n\to\infty} \Phi_n,$$

where
$$\Phi_n = \frac{2.2.4.4.6.6\ldots(2n)(2n)}{1.1.3.3.5.5\ldots(2n-1)(2n+1)}.$$

One readily shows that

$$\Phi_n = \frac{2^{4n}(n!)^4}{\{(2n)!\}^2 2n}\frac{1}{1+1/2n}.$$

From this

$$\log\Phi_n = 4n\log 2+4\log(n!)-2\log\{(2n)!\}-\log 2n-\log\frac{1}{1+1/2n}.$$

Let us substitute in this from (32). We get after reduction

$$\log\Phi_n = -2\log 2+2C+2\theta_1 B_2\frac{1}{n}-\log\frac{1}{1+1/2n}-\theta_2\frac{B_2}{2}\frac{1}{n}.$$

If in this equation we allow $n$ to become infinite, we get

$$\log\tfrac{1}{2}\pi = -2\log 2+2C.$$

From which $C = \log\sqrt{(2\pi)}$.

We thus have the final formula

$$\log(x!) = \log\sqrt{(2\pi)}+(x+\tfrac{1}{2})\log x-x+\frac{B_2}{1.2}x^{-1}+\frac{B_4}{3.4}x^{-3}+\ldots+$$

$$+\frac{B_{2k-2}}{(2k-3)(2k-2)}x^{-(2k-2)}+\theta\frac{B_{2k}}{(2k-1)(2k)}x^{-(2k-1)},$$

$$x > 1,\ 0 < \theta < 1. \quad (33)$$

This is the desired formula. Using (32) we get

$$\log(x!) = \log\sqrt{(2\pi)}+(x+\tfrac{1}{2})\log x-x+\theta\frac{B_2}{2}x^{-1},$$

$$x > 1,\ 0 < \theta < 1. \quad (34)$$

Replace $B_2$ by its value $\frac{1}{6}$ and we obtain from this the relation

$$\log\sqrt{(2\pi)}+(x+\tfrac{1}{2})\log x-x < \log(x!)$$

$$< \log\sqrt{(2\pi)}+(x+\tfrac{1}{2})\log x-x+\frac{1}{12x}. \quad (35)$$

## 12. Generalizations

We refer to § 14 of Chapter III and continue the line of reasoning initiated there. We shall use the same notation. However, in addition to the assumptions made at that point *we assume, considering $f_n(x+\omega)$ as a function of $\omega$,*

$$Pf_n(x+\omega-l)\big]_{\omega=l} = nf_{n-1}(x)$$

*and that $f_n(0) = 0$. We also note that the operators $P$ and $Q$ are each distributive with respect to addition.*

We first prove a lemma with reference to $f_n(x)$. We shall prove that

$$f_n(x+\omega)$$

$$= f_n(x)+nf_{n-1}(x)f_1(\omega)+\frac{n(n-1)}{2!}f_{n-2}(x)f_2(\omega)+\ldots+f_n(\omega). \quad (36)$$

To do this we write

$$f_n(x+\omega) = b_0(x)+b_1(x)f_1(\omega)+b_2(x)f_2(\omega)+\ldots+b_n(x)f_n(\omega), \quad (37)$$

where the $b$'s are as yet undetermined. Such an expression is possible since $f_n(x+\omega)$ is a polynomial of the $n$th degree in $\omega$. We determine the $b$'s by successively applying the operator $P$

to (37) as a function of $\omega$ and then letting $\omega = 0$ remembering that $f_n(0) = 0$.

We have attached a meaning to $f_n(x+g)$. With this meaning in mind we prove that

$$f_n\{(x+\omega)+g\} = f_n\{x+(\omega+g)\}. \tag{38}$$

Exponents are applied to $x$ and $\omega$, subscripts to $g$. Expansion of the left-hand member of (38) is the same as that given in (63) of Chapter III with $(x+\omega)$ replacing $x$. The whole expansion carried out on both sides of (38) is exactly the same as the expansions of $\{(x+\omega)+g\}^n$ and $\{x+(\omega+g)\}^n$ with the exponents replaced by the $f$-function. Thus $f_k(x)$ replaces $x^k$. Inasmuch as the binomial expansion yields an identity so does the expansion in terms of $f$ that we are considering.

Now consider any polynomial of degree $m+k$ which we write

$$\begin{aligned}\psi(x) &= a_0+a_1 f_1(x)+\ldots+a_{m+k} f_{m+k}(x),\\ \psi(x+g) &= a_0+a_1 f_1(x+g)+\ldots+a_{m+k} f_{m+k}(x+g)\\ &= a_0+a_1 F_1(x)+\ldots+a_{m+k} F_{m+k}(x).\end{aligned}$$

Hence by (62) of Chapter III

$$\begin{aligned}Q\psi(x+g) = a_k(k!)+a_{k+1}(k+1)\ldots 2f_1(x)+\ldots+\\ +a_{m+k}(m+k)(m+k-1)\ldots(m+1)f_m(x).\end{aligned}$$

That is, $$Q\psi(x+g) = P^k\psi(x), \quad m+k \geqslant 0.$$

Similarly $$Q\psi(x+\omega+g) = P^k\psi(x+\omega). \tag{39}$$

We apply Taylor's formula symbolically to this, that is, we write

$$Q\psi(x+\omega+g) = c_0(x)+c_1(x)F_1(\omega)+\ldots+c_m(x)F_m(\omega)$$

and determine the $c$'s by first letting $\omega = -g$; and then successively applying $P$, remembering (64) of Chapter III and letting $\omega = -g$ after each application.† We get

$$Q\psi(x+\omega+g) = Q\psi(x)+\sum_{\nu=1}^{m} \frac{F_\nu(\omega)}{\nu!} P^\nu Q\psi(x).$$

† Here, of course, when $\omega$ is replaced by $g$ an exponent is changed to a subscript.

Hence, by (39),

$$P^k\psi(x+\omega) = Q\psi(x)+\sum_{\nu=1}^{m}\frac{F_\nu(\omega)}{\nu!}P^\nu Q\psi(x). \tag{40}$$

This is a polynomial identity. Let $\psi(x) = P^{-k}\chi(x)$ where some particular determination is chosen in case $P^{-k}$ is not unique. Then

$$\chi(x+\omega) = \sum_{\nu=0}^{m}\frac{F_\nu(\omega)}{\nu!}P^\nu QP^{-k}\chi(x). \tag{41}$$

Like (40) this is a polynomial identity. It includes as special cases Taylor's formula, the Euler-Maclaurin formula of classical mathematics, the Boole formula, and so on.

Formula (41) will be written out in detail for two interesting special cases.

Let

$$P = \frac{d}{dx}, \qquad Q = \underset{h}{\Delta^k}, \qquad P^{-1} = \int_0^x dx, \qquad k = 1, 2,...;$$

$$\chi(x+\omega) = \underset{h}{\Delta^k}\int_0^x\int_0^{x_1}...\int_0^{x_{k-1}}\chi(x_k)\,dx_1\,dx_2...dx_k+$$

$$+B_1^{(k)}(\omega)\underset{h}{\Delta^k}\int_0^x\int_0^{x_1}...\int_0^{x_{k-2}}\chi(x_{k-1})\,dx_1\,dx_2...dx_{k-1}+...+$$

$$+\frac{1}{(k-1)!}B_{k-1}^{(k)}(\omega)\underset{h}{\Delta^k}\int_0^x\chi(x_1)\,dx_1+\frac{1}{k!}B_k^{(k)}(\omega)\underset{h}{\Delta^k}\chi(x)+...+$$

$$+\frac{1}{m!}B_m^{(k)}(\omega)\underset{h}{\Delta^k}\frac{d^{m-k}}{dx^{m-k}}\chi(x). \tag{42}$$

Secondly, we let

$$P = \Delta, \qquad Q = \frac{d^k}{dx^k},$$

$$P^{-1} = \sum_0^{x-1} \equiv \sum\Big]_x-\sum\Big]_0, \quad \text{where} \quad \sum = \Delta^{-1};$$

$$\chi(x+h) = \frac{d^k}{dx^k}\sum_{x_1=0}^{x-1}\sum_{x_2=0}^{x_1-1}\cdots\sum_{x_k=0}^{x_{k-1}-1}\chi(x_k)+$$

$$+b_1^{(k)}(\omega)\frac{d^k}{dx^k}\sum_{x_1=0}^{x-1}\sum_{x_2=0}^{x_1-1}\cdots\sum_{x_{k-1}=0}^{x_{k-2}-1}\chi(x_{k-1})+...+$$

$$+\frac{1}{(k-1)!}b_{k-1}^{(k)}(\omega)\frac{d^k}{dx^k}\sum_{x_1=0}^{x-1}\chi(x_1)+\frac{1}{k!}b_k^{(k)}(\omega)\frac{d^k}{dx^k}\chi(x)+\ldots+$$

$$+\frac{1}{m!}b_m^{(k)}(\omega)\frac{d^k}{dx^k}\Delta^{m-k}\chi(x). \tag{43}$$

In the above formulae $B_j^{(k)}(\omega)$ is the Bernoulli polynomial of order $k$ and degree $j$; similarly $b_j^{(k)}(\omega)$ is the analogous polynomial which we designate as the Bernoulli polynomial of the second kind of order $k$ and degree $j$.

Formula (42) should be compared with a related formula of Nörlund.† If the operator $\sum_{x_1=0}^{x-1}\sum_{x_2=0}^{x_1-1}\ldots\sum_{x_k=0}^{x_{k-1}-1}$ is applied to both sides of (42) we get an interesting result.

*A remainder formula.* Let us assume that $\psi(x)$ is no longer a polynomial but that it is such a function that all operations applied to it in the sequel are meaningful and valid. We consider formula (40) as our fundamental form.

We assume that

$$Q\psi(x)+\sum_{\nu=0}^{m}\{F_\nu(t)/\nu!\}P^\nu Q\psi(x)$$

involves $m+1$ points, dependent upon $x$ alone, which we call set 1. Some or all of these may coincide. For example, $\Delta^m Q\psi(x)$ involves $m+1$ distinct points and $(d^m/dx^m)Q\psi(x)$, $m+1$ coincident points. Let $x$ be a constant and $t$ a variable. When $t$ is such that $x+t$ coincides with any one of the points of set 1

$$P^k\psi(x+t)-\left[Q\psi(x)+\sum_{\nu=1}^{m}\frac{F_\nu(t)}{\nu!}P^\nu Q\psi(x)\right]$$

vanishes, inasmuch as every term is identical with the like term which is built for the polynomial of degree $m$ which coincides with $\psi(x)$ at the points of set 1. It has been remarked that (40) is an identity for a polynomial. Now denote by $p(t)$ a polynomial of degree $m+1$ which vanishes when $x+t$ coincides with any one of the $m+1$ points of set 1. Let the coefficient of $t^{m+1}$ in $p(t)$ be 1. Choose a constant $T$ such that

$$P^k\psi(x+t)-\left[Q\psi(x)+\sum_{\nu=1}^{m}\frac{F_\nu(t)}{\nu!}P^\nu Q\psi(x)\right]-Tp(t)$$

† Loc. cit., p. 160.

vanishes at the additional point where $t = w$. Then by repeated application of Rolle's theorem $(m+1)!\,T = d^{m+1}\{P^k\psi(\xi)\}/dx^{m+1}$, where $\xi$ is somewhere between the extreme values of the $(m+2)$ points composed of the points of set 1 and the additional point $x+\omega$. We have a final form for the remainder to formula (40):

$$R = \frac{1}{(m+1)!}p(\omega)\frac{d^{m+1}}{dx^{m+1}}\{P^k\psi(\xi)\}. \tag{44}$$

The corresponding formula for (41) is

$$R = \frac{1}{(m+1)!}p(\omega)\frac{d^{m+1}}{dx^{m+1}}\chi(\xi). \tag{45}$$

## EXERCISES

1. Expand $(\theta/\sin\theta)^2$ into a power series, expressing the coefficients in Bernoulli numbers.

2. Expand $\int_0^z \log(1-e^{-t})\,dt - z\log z$ into a power series and express the coefficients in terms of Bernoulli numbers.

3. Develop $\operatorname{cosec}\theta$ into a power series with coefficients expressed in terms of Bernoulli numbers.

4. Develop $\log\sin\theta$ into a power series with coefficients expressed in terms of Bernoulli numbers.

V

# STIRLING'S NUMBERS AND NUMERICAL DIFFERENTIATION

## 1. Fundamental theorems

TAYLOR'S series is as follows:

$$f(x+h) = f(x)+f'(x)\frac{h}{1!}+f''(x)\frac{h^2}{2!}+\dots. \tag{1}$$

Let $\Delta f(x) = f(x+h)-f(x)$ and $Df(x) = \dfrac{d}{dx}f(x)$. We can then write Taylor's series thus:

$$\Delta f(x) = \frac{h}{1!}Df(x)+\frac{h^2}{2!}D^2f(x)+\dots. \tag{2}$$

In other words Taylor's series serves to express the operator $\Delta$ in terms of the operator $D$. The problem of expressing $\Delta^k$ in terms of $D$ immediately presents itself.

The following formula was proved in Chapter I:

$$\Delta^k f(x) = \sum_{i=0}^{k}(-1)^i {}_kC_{k-i} f\{x+(k-i)h\}, \tag{3}$$

where ${}_kC_{k-i}$ is a binomial coefficient. Now

$$f(x+2h) = f(x)+\frac{2h}{1!}Df(x)+\frac{(2h)^2}{2!}D^2f(x)+\dots.$$

There is a similar formula for $f(x+3h),\dots.$ We can substitute these series in (3) and collect coefficients. Consequently, in a formal way what we have undertaken to do is surely possible. We wish to determine the resulting coefficients.

In order to avoid all questions of convergence we let $f(x)$ be a polynomial of degree $m$. Then

$$f(x+t) = f(x)+\frac{t}{1!}Df(x)+\frac{t^2}{2!}D^2f(x)+\dots+\frac{t^m}{m!}D^mf(x).$$

Consider $t$ as a variable and take the $k$th difference, retaining $h$ as difference interval:

$$\Delta^k f(x+t) = \frac{\Delta^k t^k}{k!}D^kf(x)+\frac{\Delta^k t^{k+1}}{(k+1)!}D^{k+1}f(x)+\dots+\frac{\Delta^k t^m}{m!}D^mf(x). \tag{4}$$

In (4) let $t = 0$:

$$\Delta^k f(x) = \frac{\Delta^k 0^k}{k!} D^k f(x) + \frac{\Delta^k 0^{k+1}}{(k+1)!} D^{k+1} f(x) + \ldots + \frac{\Delta^k 0^m}{m!} D^m f(x). \quad (5)$$

These are the desired formulae.

We let $$\mathscr{S}_j^{(k)} = \frac{1}{k!} \Delta^k 0^j, \quad h = 1.$$

Note that $$\Delta^k 0^j]_{h=h} = h^j [\Delta^k 0^j]_{h=1}.$$

We then can write (5) as follows:

$$\Delta^k f(x) = h^k \mathscr{S}_k^{(k)} D^k f(x) + \frac{h^{k+1}}{k+1} \mathscr{S}_{k+1}^{(k)} D^{k+1} f(x) + \ldots + \\ + \frac{h^m}{(k+1)\ldots m} \mathscr{S}_m^{(k)} D^m f(x). \quad (6)$$

If in (3) we replace $f(x)$ by $x^j$ and let $x = 0$, $h = 1$, we have

$$\mathscr{S}_j^{(k)} = \frac{1}{k!} \sum_{i=0}^{k-1} (-1)^i {}_kC_{k-i} (k-i)^j = \frac{1}{k!} \sum_{i=1}^{k} (-1)^{k-i} {}_kC_i \, i^j.$$

The numbers $\mathscr{S}_j^{(k)}$ are known as Stirling's numbers of the second kind.

The converse problem, namely to express $D^k$ in terms of $\Delta$, is immediately suggested. If we notice that $\mathscr{S}_k^{(k)} = 1$ and set up equation (6) for $k = 1, \ldots, m$, we have a set of equations in $Df(x), \ldots, D^m f(x)$ where the determinant of the coefficients is 1. Consequently we can always determine $Df(x), \ldots, D^m f(x)$ and thus solve our problem. However, an easier procedure is as follows. Develop $f(x+t)$ by Newton's formula. We have

$$f(x+t) = f(x) + \frac{1}{h} \frac{t^{(1)}}{1!} \Delta f(x) + \frac{1}{h^2} \frac{t^{(2)}}{2!} \Delta^2 f(x) + \ldots + \frac{1}{h^m} \frac{t^{(m)}}{m!} \Delta^m f(x), \quad (7)$$

where $t^{(j)} = t(t-h)\ldots\{t-(j-1)h\}$. Differentiate $k$ times with respect to $t$:

$$D^k f(x+t) = \frac{1}{h^k} \frac{D^k t^{(k)}}{k!} \Delta^k f(x) + \frac{1}{h^{k+1}} \frac{D^k t^{(k+1)}}{(k+1)!} \Delta^{k+1} f(x) + \ldots + \\ + \frac{1}{h^m} \frac{D^k t^{(m)}}{m!} \Delta^m f(x). \quad (8)$$

Let $t = 0$:

$$D^k f(x) = \frac{1}{h^k}\frac{D^k 0^{(k)}}{k!}\Delta^k f(x) + \frac{1}{h^{k+1}}\frac{D^k 0^{(k+1)}}{(k+1)!}\Delta^{k+1} f(x) + \ldots + \\ + \frac{1}{h^m}\frac{D^k 0^{(m)}}{m!}\Delta^m f(x). \quad (9)$$

These are the desired formulae. If we let

$$S_j^{(k)} = h^{k-j}\frac{D^k 0^{(j)}}{k!}$$

we can write

$$D^k f(x) = \frac{1}{h^k}\Big(S_k^{(k)}\Delta^k f(x) + \frac{1}{k+1}S_{k+1}^{(k)}\Delta^{k+1} f(x) + \ldots + \\ + \frac{1}{(k+1)\ldots m}S_m^{(k)}\Delta^m f(x)\Big). \quad (10)$$

The numbers $S_j^{(k)}$ are called Stirling's numbers of the first kind. It is immediate that $S_j^{(k)}$ equals the coefficient of $x^k$ in the expansion of $x(x-1)\ldots\{x-(j-1)\}$. They consequently are sometimes called *factorial coefficients*.

We now wish to extend formulae (4) and (8) to functions that are not polynomials, developing formulae for the remainder analogous to the well-known forms in Taylor's formula. We suppose $f(x)$ a function having a derivative of order $(m+1)$ at all points considered and treat (8).

Let

$$\phi(t) \equiv f(x+t) - f(x) - \frac{1}{h}\frac{t^{(1)}}{1!}\Delta f(x) - \ldots - \frac{1}{h^m}\frac{t^{(m)}}{m!}\Delta^m f(x) - Kt^{(m+1)}, \quad (11)$$

where $K$ is a constant as yet unspecified. Now $\phi(t) = 0$ when $t = 0, h, \ldots, mh$. The last term vanishes at each of these points due to the presence of $t^{(m+1)}$. The rest of the expression vanishes in as much as Newton's formula is exact at these points. Choose a positive integer $k < m+1$. By successive application of Rolle's theorem we see that $D^k\phi(t)$ vanishes at least $m+1-k$ times on the interval $0 < t < mh$. Now let $T$ be a point distinct from any of these such that $D^k T^{(m+1)} \neq 0$. Then determine $K$ so that $D^k\phi(T) = 0$. In order to be assured that $T$ be distinct from at least $m+1-k$ zeros of $D^k\phi(t)$ on the interval $0 < t < mh$, we

assume that it does not belong to this interval, although it may be an end-point. With this assumption $D^k\phi(t)$ has at least $m+2-k$ zeros on the interval delimited by $T, 0, \ldots, mh$. Hence, by repeated application of Rolle's theorem $D^{m+1}\phi(t)$ vanishes at least once on the interval delimited by $T, 0, \ldots, mh$. Hence

$$K = \frac{1}{(m+1)!} D^{m+1} f(x+\xi).$$

In particular we can choose $T = 0$ since $D^k 0^{(m+1)} \neq 0, k < m+1$. This is surely the case as $(1/k!)D^k 0^{(m+1)}$ is simply the coefficient of $t^{m+1-k}$ in the expansion of $t^{(m+1)} = t(t-h)\ldots(t-mh)$. We have the formula

$$D^k f(x+T) = \frac{1}{h^k}\frac{D^k T^{(k)}}{k!}\Delta^k f(x) + \ldots + \frac{1}{h^m}\frac{D^k T^{(m)}}{m!}\Delta^m f(x) + R_m^{(k)}, \tag{12}$$

$$R_m^{(k)} = D^{m+1} f(x+\xi)\frac{D^k T^{(m+1)}}{(m+1)!}, \tag{13}$$

where, as we have stated, $T$ does not lie within the interval $0 < t < mh$ and is so chosen that $D^k T^{(m+1)} \neq 0$, where moreover $\xi$ is within the interval delimited by $T, 0, \ldots, mh$. In particular if $T = 0$ we have

$$D^k f(x) = \frac{1}{h^k}\Big[S_k^{(k)}\Delta^k f(x) + \frac{1}{k+1}S_{k+1}^{(k)}\Delta^{k+1} f(x) + \ldots +$$

$$+\frac{1}{(k+1)\ldots m} S_m^{(k)}\Delta^m f(x)\Big] + R_m^{(k)}, \quad k < m+1, \tag{14}$$

$$R_m^{(k)} = h^{m+1-k}\frac{S_{m+1}^{(k)}}{(k+1)\ldots(m+1)} D^{m+1} f(x+\xi), \quad T = 0. \tag{15}$$

Now let us consider (4) and suppose that $f(x)$ is a function such that $D^{m+1}f(x)$ exists at all points considered. Form the function

$$\psi(t) = f(x+t) - f(x) - \frac{t}{1!}Df(x) - \frac{t^2}{2!}D^2 f(x) - \ldots - \frac{t^m}{m!}D^m f(x) - Kt^m,$$

where $K$ is a constant as yet unspecified. This function has a zero of order $(m+1)$ when $t = 0$. Consequently $D^k\psi(t)$ has a zero of order $(m+1-k)$ when $t = 0$. Now let $T$ be chosen such that $\Delta^k T^{m+1} \neq 0$. Then choose $K$ so that $\Delta^k\psi(t)$ vanishes when $t = T$.

But by the mean-value theorem $\Delta^k\psi(T) = h^k D^k\psi(\delta)$, where $T < \delta < T+kh$. Consequently $D^k\psi(\delta) = 0$, where

$$T < \delta < T+kh.$$

We additionally restrict $T$ so that this interval does not include the origin, although the origin may be an end-point. Under these restrictions $D^k\psi(t)$ vanishes $m+2-k$ times in the interval delimited by $0$, $T$, $T+kh$ with the end-points excluded. Hence $D^{m+1}\psi(t)$ vanishes at least once on this interval. Hence

$$K = \frac{1}{(m+1)!} D^{m+1}f(x+\xi),$$

where $\xi$ is on the interval delimited by $0$, $T$, $T+kh$. It is well to remark that $T$ can have any positive value. This is true since $\Delta^k T^{m+1} = h^k D^k \eta^{m+1} \neq 0$ because $\eta > T > 0$ and $k < m+1$.

We consequently have the formula

$$\Delta^k f(x+T) = \frac{\Delta^k T^k}{k!} D^k f(x) + \ldots + \frac{\Delta^k T^m}{m!} D^m f(x) + \mathscr{R}_m^{(k)}, \tag{16}$$

$$\mathscr{R}_m^{(k)} = D^{m+1}f(x+\xi)\frac{\Delta^k T^{m+1}}{(m+1)!}, \tag{17}$$

where $T$ is restricted as above delineated. If $T = 0$ we have

$$\Delta^k f(x) = h^k \mathscr{S}_k^{(k)} D^k f(x) + \frac{h^{k+1}\mathscr{S}_{k+1}^{(k)}}{k+1} D^{k+1}f(x) + \ldots + \frac{h^m \mathscr{S}_m^{(k)}}{(k+1)\ldots m} D^m f(x) + \mathscr{R}_m^{(k)}, \tag{18}$$

$$\mathscr{R}_m^{(k)} = \frac{h^{m+1}\mathscr{S}_{m+1}^{(k)}}{(k+1)\ldots(m+1)} D^{m+1}f(x+\xi), \quad T = 0. \tag{19}$$

Remainder terms as given in (15) and (17) are known as Markoff's forms.

Formulae for the remainder which have been obtained are analogous to the Lagrange form for the remainder in Taylor's formula. The problem of obtaining an integral (Cauchy) form for the remainder suggests itself. The following formula follows from (3) by the subtraction of $\sum_{i=0}^{k} (-1)^i {}_kC_{k-i} f(x) = (1-1)^k f(x) = 0$:

$$\Delta^k f(x) = \sum_{i=0}^{k-1} (-1)^i {}_kC_{k-i}[f\{x+(k-i)h\} - f(x)]. \tag{20}$$

Now in (20) expand $f\{x+(k-i)h\}-f(x)$ by Taylor's formula using the integral form for the remainder. Denote the sum of these remainders by $\mathscr{R}_m^{(k)}$. Then

$$\mathscr{R}_m^{(k)} = \sum_{i=0}^{k-1} (-1)^i {}_kC_{k-i} \frac{1}{m!} \int\limits_x^{x+(k-i)h} \{x+(k-i)h-t\}^m D^{m+1}f(t)\,dt$$
$$= \frac{1}{m!} \Delta_z^k \int\limits_x^z (z-t)^m D^{m+1}f(t)\,dt\Big]_{z=x}. \tag{21}$$

Here $\Delta_z^k$ means the difference with $z$ as variable.

The calculation of $D^kf(x)$ by means of formula (14) is known as *numerical differentiation.*

## 2. Operational methods

We now shall start over again so to speak. Let us call attention to formula (2). This can be symbolically written

$$\Delta f(x) = (e^{hD}-1)f(x).$$

We write symbolically $\Delta = (e^{hD}-1)$. If the Taylor's series converges this symbolic interpretation can be given a genuine significance. If $f(x)$ is a polynomial there is no question as to the interpretation of the operator $(e^{hD}-1)$ nor as to the fact that its successive application obeys all the laws of ordinary multiplication. More specifically:

Assume $f(x)$ of degree $m$. All terms in the formal expansion of $(e^{hD}-1)$ of higher power in $D$ than $m$ will yield zero when operating on $f(x)$. We then in fact are dealing only with

$$hD+\frac{h^2D^2}{2!}+\ldots+\frac{h^mD^m}{m!},$$

an operator of a finite number of terms. We write

$$\Delta^k f(x) = (e^{hD}-1)^k f(x). \tag{22}$$

Equation (22) is valid for the polynomial and is an abbreviated way to write (5). But if $k > 0$,

$$(e^x-1)^k = \left(x+\frac{x^2}{2!}+\ldots+\frac{x^m}{m!}\right)^k + \text{(terms of degree higher than } m).$$

Since $m$ is any positive integer, by comparing (6) and (22) we have

$$(e^x-1)^k = \mathscr{S}_k^{(k)}x^k+\frac{1}{k+1}\mathscr{S}_{k+1}^{(k)}x^{k+1}+\frac{1}{(k+1)(k+2)}\mathscr{S}_{k+2}^{(k)}x^{k+2}+\dots. \tag{23}$$

We now consider the inverse problem. We note that

$$S_j^{(1)} = D0^j = (-1)^{j-1}\{(j-1)!\}.$$

Consequently by (10)

$$Df(x) = \frac{1}{h}[\Delta f(x)-\tfrac{1}{2}\Delta^2 f(x)+\tfrac{1}{3}\Delta^3 f(x)-\dots],$$

an equation surely valid if $f(x)$ is a polynomial. Symbolically we write

$$Df(x) = \frac{1}{h}[\log(1+\Delta)]f(x) \tag{24}$$

or
$$D = \frac{1}{h}\log(1+\Delta).$$

Exactly as in the previous case

$$D^m f(x) = \frac{1}{h^m}[\log(1+\Delta)]^m f(x).$$

This is a symbolic method of writing (10).

The function $(1/k!)(e^x-1)^k$ is called a *generating function* for Stirling's numbers of the second kind of order $k$. Its development into a power series affords a method of calculating these numbers. Similarly $(1/k!)\{\log(1+x)^k\}$ is a generating function for Stirling's numbers of the first kind of order $k$.

## 3. Analogues and generalizations of Stirling's numbers

To get $\Delta^k$ in terms of $D$ we took Taylor's formula

$$f(x+t) = f(x)+tDf(x)+\frac{t^2}{2!}D^2f(x)+\dots+\frac{t^m}{m!}D^m f(x)$$

and applied $\Delta^k$ operating on $t$ to both sides and let $t = 0$. There is nothing peculiar about $\Delta^k$ as an operator. We can apply $F(\Delta)$, where $F$ is a polynomial. We can take irregularly spaced points and apply the divided difference operator, the operator of the

mean, etc. However, we shall make a brief discussion of an interesting operator or two.

Operate on both sides of Taylor's formula above with $h\sum$ where $h\sum = \underset{h}{\Delta}^{-1}$. This is not a unique operator. When we operate within the domain of polynomials, as we shall, there is an additive constant as in integration. Let

$$h\sum f(x+t) = \phi(x+t).$$

Note moreover that $h\sum nx^{n-1} = B_n(x, h)+C$ and that

$$f(x) = \underset{h}{\Delta}\phi(x).$$

We get

$$h\sum f(x+t) = B_1(t,h)\underset{h}{\Delta}\phi(x)+\frac{B_2(t,h)}{2!}\underset{h}{\Delta}D\phi(x)+$$

$$+\frac{B_3(t,h)}{3!}\underset{h}{\Delta}D^2\phi(x)+\ldots+\frac{B_{m+1}(t,h)}{(m+1)!}\underset{h}{\Delta}D^m\phi(x)+C = \phi(x+t).$$

We know from an analogue of Chapter III (10), where $h \neq 1$, that

$$\int_0^h B_j(t,h)\,dt = 0, \quad j > 0.$$

As a result

$$C = \frac{1}{h}\int_0^h \phi(x+t)\,dt = \frac{1}{h}\int_x^{x+h} \phi(t)\,dt = \underset{h}{\Delta}\int_0^x \phi(t)\,dt.$$

We consequently write

$$h\sum f(x+t) = \phi(x+t) = \underset{h}{\Delta}\int_0^x \phi(t)\,dt+B_1(t,h)\underset{h}{\Delta}\phi(x)+$$

$$+\frac{B_2(t,h)}{2!}\underset{h}{\Delta}D\phi(x)+\ldots+\frac{B_{m+1}(t,h)}{(m+1)!}\underset{h}{\Delta}D^m\phi(x).$$

Let $t = 0$, and this expresses $\sum$ in terms of $D$, $D^2$,... and $\underset{h}{\Delta}$. The Bernoulli numbers appear as analogues of the Stirling numbers. In case $h = 1$ we have the classical Bernoulli polynomials and numbers of Chapter III and the classical Euler-Maclaurin formula. Repeated application of $h\sum$ yields (42) of Chapter IV. As a matter of fact an even more general

formula can be obtained if $h$ is changed every time the operator $h\sum$ is applied.

An interesting operator is $\nabla^{-1}$, where $\nabla f(x) = \dfrac{f(x+h)+f(x)}{2}$, the operator of the mean. There is no lack of uniqueness and we are immediately led to the Boole formula,

$$h\sum f(x+t) = \phi(x+t) = \sum_{\nu=0}^{m-1} \frac{E_\nu(t)}{\nu!}\nabla D^\nu \phi(x),$$

where $E_\nu(t)$ is the Euler polynomial of order $\nu$.

This may be run to any order by repeated application of $M^{-1}$. Let $t = 0$ and $\sum f(x)$ is expressed in terms of $D$, $D^2$,... and $\nabla$ with the Euler numbers replacing the Stirling numbers.

But why begin with Taylor's formula? Try any of the other formulae for $f(x+t)$ which we have discussed. Newton's formula, for example, is the simplest:

$$f(x+t) = f(x)+\frac{t}{h}\Delta f(x)+\frac{t^{(2)}}{2!\,h^2}\Delta^2 f(x)+\ldots+\frac{t^{(m)}}{m!\,h^m}\Delta^m f(x).$$

Let us apply $\int dt$ to this $k$ times. We are led to (43) of Chapter IV. If we apply the inverse mean $M^{-1}$ to Newton's formula we get a new set of polynomials and a new formula which is not without interest.

Let us begin not with Taylor's formula nor with Newton's formula nor with the Euler-Maclaurin formula but with a general formula that includes all of these. We have the following polynomial formula, for which see Chapter IV (41),

$$\chi(x+t) = \sum_{\nu=0}^{m} \frac{F_\nu(t)}{\nu!} P^\nu Q P^{-k}\chi(x), \tag{25}$$

where $Q$ is a second operator more specifically determining the polynomials $F_n(x)$. Now on any one of these formulae as a function of $t$ we can operate with any linear operator $V$, getting other formulae; and in case $V$ is such that

$$Vf(x+t-e)\big|_{t=e} = Vf(x)$$

we express $V$ in terms of the operators $P$ and $Q$ of the formula. The Stirling numbers are replaced by $VF_n(t)\big|_{t=0}$. The factorials used in the classical definition are trivial.

*Stirling's Numbers of the First Kind:* $S_n^{(m)}$

| $n \backslash m$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | 1 | | | | |
| 2 | −1 | 1 | | | |
| 3 | 2 | −3 | 1 | | |
| 4 | −6 | 11 | −6 | 1 | |
| 5 | 24 | −50 | 35 | −10 | 1 |
| 6 | −120 | 274 | −225 | 85 | −15 |
| 7 | 720 | −1764 | 1624 | −735 | 175 |
| 8 | −5040 | 13068 | −13132 | 6769 | −1960 |
| 9 | 40320 | −109584 | 118124 | −67284 | 22449 |
| 10 | −362880 | 1026576 | −1172700 | 723680 | −269325 |
| 11 | 3628800 | −10628640 | 12753576 | −8409500 | 3416930 |
| 12 | −39916800 | 120543840 | −150917976 | 105258076 | −45995730 |

*Stirling's Numbers of the Second Kind:* $\mathscr{S}_n^{(m)}$

| $n \backslash m$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1 | | | | | |
| 2 | 1 | 1 | | | | |
| 3 | 1 | 3 | 1 | | | |
| 4 | 1 | 7 | 6 | 1 | | |
| 5 | 1 | 15 | 25 | 10 | 1 | |
| 6 | 1 | 31 | 90 | 65 | 15 | 1 |
| 7 | 1 | 63 | 301 | 350 | 140 | 21 |
| 8 | 1 | 127 | 966 | 1701 | 1050 | 266 |
| 9 | 1 | 255 | 3025 | 7770 | 6951 | 2646 |
| 10 | 1 | 511 | 9330 | 34105 | 42525 | 22827 |
| 11 | 1 | 1023 | 28501 | 145750 | 246730 | 179487 |
| 12 | 1 | 2047 | 86526 | 611501 | 1379400 | 1323652 |

## EXERCISES

1. Show that
$$\mathscr{S}_0^{(m)} = S_0^{(m)} = 0, \quad m \neq 0, \quad \text{and} \quad \mathscr{S}_0^{(0)} = S_0^{(0)} = 1.$$

2. Prove $$\mathscr{S}_{j+1}^{(m)} = m\mathscr{S}_j^{(m)} + \mathscr{S}_j^{(m-1)}.$$

3. Prove $$S_{j+1}^{(m)} = S_j^{(m-1)} - jS_j^{(m)}.$$

4. Prove $$\mathscr{S}_{n+1}^{(n)} = \frac{n(n+1)}{2}\mathscr{S}_n^{(n)}.$$

5. Prove
$$n\mathscr{S}_{n-1}^{(m)} + \frac{n(n-1)}{2!}\mathscr{S}_{n-2}^{(m)} + \dots + \frac{n!}{(n-m)!}\mathscr{S}_{n-m}^{(m)} = (m+1)\mathscr{S}_n^{(m+1)}.$$

6. Show that
$$S_n^{(1)} + S_n^{(2)} + S_n^{(3)} + \dots + S_n^{(n)} = 0, \quad n > 1.$$

7. Prove the symbolic equation
$$\log(1+\Delta)0^y = 0, \quad y > 1.$$

8. Expand $x^{(m)}x^{(n)}$ by Taylor's formula. Express the result by means of Stirling's numbers.

9. Calculate by the use of formula (14)
$$D^5 \tan x]_{x=\frac{1}{8}\pi},$$
using a trigonometric table of tangents. Discuss the accuracy of your result.

10. Calculate by means of formula (14)
$$D^5 \cot x]_{x=40^\circ,\ h=1^\circ},$$
using tabular values for cot 40°, cot 41°,.... Discuss the accuracy of your result.

11. Calculate $\Delta^5 \sin 40^\circ$, $h = 1^\circ$, by the use of formula (18). Discuss the accuracy of your result.

12. Calculate $D^3 \sin 40^\circ\, 30'$, using formula (14) and tabular values for the functions of 40°, 41°,....

13. Generalize formula (42) of Chapter IV by changing $h$ each time the operator $\Delta$ is applied.

14. Derive an analogous formula to (42) of Chapter IV, where $D$ replaces $\Delta$ and $\sum$ replaces $\int dx$.

15. Express the Bernoulli numbers in terms of Stirling's numbers.

16. Express the Bernoulli numbers of the second kind in terms of Stirling's numbers.

# VI

# INTERPOLATION AND MECHANICAL QUADRATURES

## A. INTERPOLATION

### 1. Statement of the problem

THE general problem of interpolation can be formulated as follows:

*A function $f(x)$ is known for certain distinct values of the argument $x_1, x_2,..., x_n$:*

$$f(x_1) = y_1, \quad f(x_2) = y_2, \quad ..., \quad f(x_n) = y_n. \tag{1}$$

*Find its approximate value for another given value of the argument.* We use the value of a function $F(x)$, which coincides with $f(x)$ at the given points $x_1, x_2,..., x_n$ and which is given by a simple formula. This usually gives us the approximate value desired. The function $F(x)$ is called an *interpolation function*:

$$F(x_i) = f(x_i) = y_i, \quad i = 1, 2,..., n. \tag{2}$$

We then write the approximate equality

$$f(x) \simeq F(x) \tag{3}$$

($\simeq$ means approximately equal).

The importance of interpolation arises from the following fact. If the 'interpolating function' $F(x)$ is chosen so that there is but small error throughout $(a, b)$, then in many considerations $f(x)$ can be replaced by the simpler function $F(x)$. Of particular importance in applications is the following approximate equality to be discussed later:

$$\int_a^b f(x)\,dx \simeq \int_a^b F(x)\,dx. \tag{4}$$

Two factors play an important role in the theory of interpolation: (i) construction of the interpolating function; (ii) estimate of the 'remainder' or 'error', that is, of the difference $f(x)-F(x)$ in (3). It is usual to choose for $F(x)$ a polynomial or a trigonometric sum. We thus speak respectively of 'polynomial' or 'trigonometric' interpolation. Our present discussion

is confined to polynomial interpolation. Nothing can be said about the remainder in (3) if no further assumptions are made concerning the nature of the function $f(x)$, for the data (1) alone do not determine the behaviour of $f(x)$ elsewhere in $(a, b)$. As a matter of fact we shall restrict the class of functions to be studied to those having continuous derivatives in $(a, b)$, up to a prescribed order. We may then enlarge the data (1) so as to include preassigned values of various derivatives of $f(x)$.

## 2. Lagrange interpolation formula

We consider first the simplest case where the values of $f(x)$ are preassigned at $n$ given points, all distinct, and where no derivatives are involved. We seek to construct a polynomial of as low a degree as possible which will coincide with $f(x)$ at the given points. The $n$ points impose $n$ conditions, and we expect in general to require a polynomial of degree $n-1$, which involves $n$ coefficients.

Let the given points be $x_1, x_2,..., x_n$ and let

$$\omega(x) = (x-x_1)(x-x_2)...(x-x_n), \tag{5}$$

$$\begin{aligned} \omega_1(x) &= \frac{\omega(x)}{(x-x_1)\omega'(x_1)} = \frac{(x-x_2)(x-x_3)...(x-x_n)}{(x_1-x_2)(x_1-x_3)...(x_1-x_n)}, \\ &\cdots\cdots\cdots\cdots\cdots \\ \omega_n(x) &= \frac{\omega(x)}{(x-x_n)\omega'(x_n)} = \frac{(x-x_1)(x-x_2)...(x-x_{n-1})}{(x_n-x_1)(x_n-x_2)...(x_n-x_{n-1})}. \end{aligned} \tag{6}$$

These $\omega_i(x)$, $i = 1, 2,.., n$, are called 'fundamental Lagrangian functions'. They are polynomials of degree $n-1$, with the following characteristic property, as seen from (6):

$$\omega_i(x_j) = \begin{cases} 0, & j \neq i, \\ 1, & j = i. \end{cases} \tag{7}$$

It follows that

$$L_{n-1}(x) \equiv \sum_{i=1}^{n} \omega_i(x) f(x_i) = \sum_{i=1}^{n} \frac{\omega(x)}{(x-x_i)\omega'(x_i)} f(x_i) \tag{8}$$

is an interpolation polynomial. Its degree does not exceed $n-1$ and, by (7),

$$L_{n-1}(x_i) = f(x_i), \quad i = 1, 2,..., n. \tag{9}$$

*The polynomial* $L_{n-1}(x)$ *is known as the Lagrange interpolation polynomial.* Moreover, it is the *only interpolation polynomial, of degree not greater than* $n-1$, *because the data of* (9) *determine such a polynomial* uniquely. In fact, let a second polynomial, $P(x)$, of degree not greater than $n-1$, be such that $P(x_i) = f(x_i)$, $i = 1, 2,..., n$. Then $L_{n-1}(x)-P(x)$ is of degree not greater than $n-1$ and vanishes at $n$ points. This is impossible unless all of its coefficients are zero.

We now pass to the remainder in the approximate equality (3) which we rewrite as

$$f(x) = L_{n-1}(x)+R_n(x). \tag{10}$$

We call this the Lagrange interpolation formula with remainder.

Let $x$ be different from $x_1,..., x_n$ and fixed. Introduce the function

$$\phi(z) \equiv f(z)-L_{n-1}(z)-K\omega(z),$$

where $K$ is a constant so chosen that

$$\phi(x) = f(x)-L_{n-1}(x)-K\omega(x) = 0. \tag{11}$$

By (5), (9), and (11) $\phi(x)$ has $n+1$ roots: $x_1$, $x_2,...,$ $x_n$ and $x$. By repeated application of Rolle's theorem $\phi^{(n)}(x)$ has at least one root lying between the smallest and the largest of the numbers $x$, $x_1,...,$ $x_n$. That is,

$$\phi^{(n)}(\xi) = 0.$$

But $$L_{n-1}^{(n)}(z) \equiv 0 \quad \text{and} \quad \omega^{(n)}(\xi) = n!.$$

Consequently $$K = \frac{f^n(\xi)}{n!}.$$

Equation (11) now yields the formula

$$f(x) = L_{n-1}(x)+\frac{f^n(\xi)}{n!}\omega(x). \tag{12}$$

Hence the remainder in the approximate formula (3) is

$$R_n(x) = \frac{f^{(n)}(\xi)}{n!}(x-x_1)(x-x_2)...(x-x_n). \tag{13}$$

Here and hereafter in this chapter $\xi$ denotes a number lying between the smallest and the largest of the numbers $x$, $x_1,...,$ $x_n$, different in different formulae.

Formula (13) shows that the Lagrange interpolation formula is *exact*, if $f(x)$ is a polynomial of degree less than or equal to $n-1$. Furthermore, it enables us to estimate the error in (3). Assume that all $x$'s lie between $a$ and $b$. So therefore does $\xi$. Let

$$M_n = \max_{a\leqslant x\leqslant b} |f^{(n)}(x)|$$

and

$$\Omega_n = \max_{a\leqslant x\leqslant b} |\omega(x)|.$$

Then

$$|R_n(x)| \leqslant \frac{M_n}{n!}\Omega_n, \tag{14}$$

$$|R_n(x)| \leqslant \frac{M_n(b-a)^n}{n!}, \quad a \leqslant x \leqslant b. \tag{15}$$

In practice, the functional values $f(x_i)$ may have been obtained from observations, and consequently be affected by certain observational errors $\pm\epsilon_i$, so that instead of (8) we have

$$L_{n-1}(x)+\Delta L_{n-1}(x) = \sum_{i=1}^{n} \omega_i(x)[f(x_i)\pm\epsilon_i], \tag{16}$$

$$\Delta L_{n-1}(x) = \sum_{i=1}^{n} \pm\epsilon_i\, \omega_i(x).$$

This shows the effect of the errors $\pm\epsilon_i$ on the polynomial $L_{n-1}(x)$. Moreover, if $|\epsilon_i| \leqslant \epsilon$, for $i = 1, 2,..., n$, then

$$|\Delta L_{n-1}(x)| \leqslant \epsilon \sum_{i=1}^{n} |\omega_i(x)|.$$

Here the last factor may become large at some points, as $n$ grows. So even small errors in the functional values $f(x_i)$ may produce a considerable change, $\Delta L_{n-1}(x)$. The possibility that the Lagrange interpolation formula thus aggravate errors is a serious disadvantage, especially when dealing with empirical functions.

Another disadvantage of (8) lies in its structure. If we add to our data new $x_i, f(x_i)$, we must start the computation of the new interpolation polynomial all over again, and the previous labour is lost. In order to avoid this, we write the interpolation polynomial in a new form as follows:

$$L_{n-1}(x) = l_0+l_1(x-x_1)+l_2(x-x_1)(x-x_2)+...+ \\ +l_n(x-x_1)...(x-x_n), \tag{17}$$

and seek to determine the coefficients $l_0$, $l_1,\ldots,$ $l_{n-1}$ so as to satisfy (9). We find successively, introducing divided differences (see Chapter I),

$$L_{n-1}(x) = f(x_1)+[x_2\,x_1](x-x_1)+[x_3\,x_2\,x_1](x-x_1)(x-x_2)+\ldots+$$
$$+[x_n\,x_{n-1}\ldots x_1](x-x_1)(x-x_2)\ldots(x-x_n).$$

Thus

$$f(x) = f(x_1)+[x_2\,x_1](x-x_1)+[x_3\,x_2\,x_1](x-x_1)(x-x_2)+\ldots+$$
$$+[x_n\,x_{n-1}\ldots x_1](x-x_1)(x-x_2)\ldots(x-x_n)+R_n. \quad (18)$$

The remainder is again given by (13). Formula (18) is called *Newton's interpolation formula with divided differences.* A special case, important in applications, is where the $x_i$ are equidistant:

$$x_1 = a, \quad x_2 = a+h, \quad \ldots, \quad x_n = a+(n-1)h,$$
$$h = \frac{b-a}{n-1}. \quad (19)$$

Formula (18) now becomes *Newton's interpolation formula:*

$$f(x) = L_{n-1}(x)+R_n(x)$$
$$= f(a)+\frac{\Delta f(a)}{1!\,h}(x-a)^{(1)}+\frac{\Delta^2 f(a)}{2!\,h^2}(x-a)^{(2)}+\ldots+$$
$$+\frac{\Delta^{n-1}f(a)}{(n-1)!\,h^{n-1}}(x-a)^{(n-1)}+R_n(x), \quad (20)$$
$$R_n(x) = \frac{(x-a)^{(n)}f^{(n)}(\xi)}{n!}.$$

Note the similarity of (20) to Taylor's formula. Taking $n = 2$ in (20), we get

$$f(x) = f(a)+\frac{f(b)-f(a)}{b-a}(x-a)+R_2(x), \quad (21)$$
$$R_2(x) = \frac{f''(\xi)}{2}(x-a)(x-b);$$

$\xi$ lying between the largest and the smallest of the numbers $x$, $a$, $b$. This is the so-called linear interpolation formula or the formula for interpolation by proportional parts. In geometrical language we replace the arc of the curve from $x = a$ to $x = b$ by a straight line. An advantage in (21) lies in the fact that the

factor multiplying $f''(\xi)$ in $R_2(x)$, namely, $(x-a)(x-b)$ keeps a constant sign when $a < x < b$. This is of aid in estimating $R_2(x)$. Thus, if the sign of $f''(x)$ does not change in $(a, b)$, $R_2(x)$ is of constant opposite sign. Moreover, $m \leqslant f''(x) \leqslant M$ implies

$$\frac{(x-a)(x-b)}{2}M \leqslant R_2(x) \leqslant \frac{(x-a)(x-b)}{2}m \tag{22}$$

when $a \leqslant x \leqslant b$. Also $|f''(x)| < A$ implies

$$|R_2(x)| \leqslant \frac{(b-a)^2}{8}A, \tag{23}$$

since

$$0 \leqslant (x-a)(b-x) = \left(\frac{b-a}{2}\right)^2 - \left(\frac{b+a}{2}-x\right)^2 \leqslant \left(\frac{b-a}{2}\right)^2, \qquad a \leqslant x \leqslant b.$$

*Illustrations*:

(i) Given $f(0) = 1$, $f(-1) = 2$, $f(1) = 1$ construct the corresponding interpolation polynomial.

Here $n = 3$. By (8),

$$L_2(x) = \frac{(x+1)(x-1)}{(0+1)(0-1)}.1+\frac{(x-0)(x-1)}{(-1-0)(-1-1)}.2+\frac{(x-0)(x+1)}{(1-0)(1+1)}.1$$
$$= \frac{x^2}{2}-\frac{x}{2}+1.$$

By (18), Newton's formula with divided differences,

$$L_2(x) = 1+\frac{2-1}{-1-0}(x-0)+\frac{\dfrac{1-2}{1+1}-\dfrac{2-1}{-1-0}}{1-0}(x-0)(x+1)$$
$$= \frac{x^2}{2}-\frac{x}{2}+1.$$

In order to make use of Newton's formula (20) we arrange our data taking $x_1 = -1$, $x_2 = 0$, $x_3 = 1$. Then $h = 1$, $a = -1$, $\Delta f(-1) = -1$, $\Delta^2 f(-1) = 1$, and hence

$$L_2(x) = 2+\frac{-1}{1!}(x+1)+\frac{1}{2!}(x+1)x = \frac{x^2}{2}-\frac{x}{2}+1.$$

(ii) As a second illustration, we apply linear interpolation to the computation of logarithms and antilogarithms. This will

show that even the simplest linear interpolation formula (21) is capable of yielding good numerical results if the second derivative of the function in question is small in the interval considered.

Given a five-place table of common logarithms of integers, $N$ from 1000 to 9999 inclusive, we wish to compute $\log(N+x)$ from the tabular values of $\log N$ and $\log(N+1)$, $0 < x < 1$. We use proportional parts; that is, formula (21). We get

$$\log(N+x) = \log N + x[\log(N+1) - \log N] + R_2(x) \tag{24}$$

with an error

$$R_2(x) = \frac{-x(x-1)}{2} \frac{\log e}{(N+\xi)^2} > 0, \quad 0 < \xi < 1. \tag{25}$$

Now $e < 3 < \sqrt{10}$, $\log e < \frac{1}{2}$, $\frac{1}{N+\xi} < \frac{1}{N} \leqslant 10^{-3}$, and $(1-x)x < \frac{1}{4}$. Consequently,

$$R_2(x) < \frac{1}{16} 10^{-6} < 7.10^{-8}.$$

Thus, the error has no influence on the seventh digit of $\log(N+x)$, as furnished by our approximate formula.

The computation of an antilogarithm consists in finding $x$, $0 < x < 1$, from the known values of $\log(N+x)$, $\log N$, $\log(N+1)$. We can determine $x$ from the same formula, namely (24). Solving for $x$, we find

$$x = \frac{\log(N+x) - \log N}{\log(N+1) - \log N} + \bar{R}_2(x)$$

with an error, which by (25) is given as follows:

$$\bar{R}_2(x) = -\frac{R_2(x)}{\log(N+1) - \log N}$$

$$= \frac{x(x-1)}{2[\log(N+1) - \log N]} \frac{\log e}{(N+\xi)^2} < 0,$$

$$|\bar{R}_2(x)| < \frac{1}{8} \frac{\log e}{\log(N+1) - \log N} \frac{1}{(N+\xi)^2}.$$

We have further†

$$\log(1+z) > \frac{z}{h+1} \log e, \quad 0 < z \leqslant h+1.$$

† This is readily established by applying the mean-value theorem to $\log(1+z)$.

Hence, since

$$\frac{1}{N} \leqslant 10^{-3} \quad \text{and} \quad \frac{N}{N+\xi} < 1,$$

$$|\bar{R}_2(x)| < \frac{1}{8}(10^{-3}+1)\frac{N}{(N+\xi)^2} < \frac{1}{8N}\,\frac{10^3+1}{10^3} < \frac{1}{4}10^{-3}.$$

## 3. A general interpolation formula involving derivatives

We wish to construct a polynomial $F(x)$ of lowest possible degree such that

$$\begin{array}{llll}
F(x_1) = f(x_1), & F'(x_1) = f'(x_1), & \ldots, & F^{(\alpha_1-1)}(x_1) = f^{(\alpha_1-1)}(x_1), \\
F(x_2) = f(x_2), & F'(x_2) = f'(x_2), & \ldots, & F^{(\alpha_2-1)}(x_2) = f^{(\alpha_2-1)}(x_2), \\
\cdot & \cdot & \cdot & \cdot \\
F(x_m) = f(x_m), & F'(x_m) = f'(x_m), & \ldots, & F^{(\alpha_m-1)}(x_m) = f^{(\alpha_m-1)}(x_m).
\end{array} \tag{26}$$

Here the points $x_1,\ldots, x_m$ are distinct. The total number of data is

$$\alpha_1+\alpha_2+\ldots+\alpha_m = n \tag{27}$$

and we suspect that in general the degree of $F(x)$ will be $n-1$ and that this $F(x)$ will be unique. In its construction we make use in succession of the first row of data in (26), then the second row, and so on. We write

$$\begin{aligned}
F(x) = {} & l_{10}+l_{11}(x-x_1)+l_{12}(x-x_1)^2+\ldots+l_{1,\alpha_1-1}(x-x_1)^{\alpha_1-1} \\
& +(x-x_1)^{\alpha_1}[l_{20}+l_{21}(x-x_2)+l_{22}(x-x_2)^2+\ldots+l_{2,\alpha_2-1}(x-x_2)^{\alpha_2-1}]+ \\
& +(x-x_1)^{\alpha_1}(x-x_2)^{\alpha_2}[l_{30}+l_{31}(x-x_3)+l_{32}(x-x_3)^2+\ldots+ \\
& \qquad\qquad +l_{3,\alpha_3-1}(x-x_3)^{\alpha_3-1}]+ \\
& \cdots\cdots\cdots\cdots\cdots\cdots \\
& +(x-x_1)^{\alpha_1}(x-x_2)^{\alpha_2}\ldots(x-x_{m-1})^{\alpha_{m-1}}[l_{m0}+l_{m1}(x-x_m)+ \\
& \qquad\qquad +l_{m2}(x-x_m)^2+\ldots+l_{m,\alpha_m-1}(x-x_m)^{\alpha_m-1}].
\end{aligned} \tag{28}$$

The coefficients $l_{ij}$ are determined by rows, starting with the first. The bracketed polynomial in the second row vanishes with all derivatives up to the order $\alpha_1-1$ inclusive when $x = x_1$. The bracketed polynomial in the third row vanishes with all

derivatives up to the order $\alpha_1+\alpha_2-1$ inclusive at the points $x_1$, $x_2$, and so on. Thus

$$l_{10} = F(x_1) = f(x_1), \quad l_{11} = \frac{F'(x_1)}{1!} = \frac{f'(x_1)}{1!}, \quad \ldots,$$

$$l_{1,\alpha_1-1} = \frac{f^{\alpha_1-1}(x_1)}{(\alpha_1-1)!} \tag{29}$$

and the polynomial in the first row in (28) is completely determined. Denote it by $F_1(x)$. Now

$$l_{20}+l_{21}(x-x_2)+\ldots+l_{2,\alpha_2-1}(x-x_2)^{\alpha_2-1} = \frac{F(x)-F_1(x)}{(x-x_1)^{\alpha_1}}+(x-x_2)^{\alpha_2}P(x).$$

Here $P(x)$ is a polynomial. We find, as in (29),

$$l_{20} = \frac{f(x)-F_1(x)}{(x-x_1)^{\alpha_1}}\bigg|_{x=x_2} = \frac{f(x_2)-F_1(x_2)}{(x_2-x_1)^{\alpha_1}},$$

$$l_{21} = \frac{d}{dx}\left[\frac{f(x)-F_1(x)}{(x-x_1)^{\alpha_1}}\right]_{x=x_2}, \quad \ldots, \tag{30}$$

$$l_{2,\alpha_2-1} = \frac{d^{\alpha_2-1}}{dx^{\alpha_2-1}}\left[\frac{f(x)-F_1(x)}{(x-x_1)^{\alpha_1}}\right]_{x=x_2},$$

and so on with the third row, fourth row,.... By such arrangement, adding a new row of data in (26) means increasing (28) by a new row, without invalidating the calculations already performed.

We next consider the remainder where $F(x)$ is given by (28); we proceed as follows: Introduce the function

$$\phi(z) = f(z)-F(z)-K(z-x_1)^{\alpha_1}(z-x_2)^{\alpha_2}\ldots(z-x_m)^{\alpha_m},$$

where the constant $K$ is so chosen that $\phi(z) = 0$ when $z = x$. The function $\phi(z)$ now has $n+1$ zeros on the interval delimited by $x$, $x_1$,..., $x_m$. By a repeated application of Rolle's theorem, we arrive at the following result:

$$f(x) = F(x)+R_n(x), \tag{31}$$

$$R_n(x) = \frac{f^n(\xi)}{n!}(x-x_1)^{\alpha_1}(x-x_2)^{\alpha_2}\ldots(x-x_m)^{\alpha_m},$$

where $\xi$ lies on the interval delimited by $x_1,\ldots, x_m$ and $x$. Moreover, $R_n(x)$ is a continuous function of $x$ since $f(x)$ and $F(x)$ are continuous. It results that $f^n(\xi)$ is continuous in $x$ so long as

$$x \neq x_1, x_2,\ldots, x_n.$$

The special case $m = 1$ yields Taylor's formula with the classical Lagrange's remainder. Thus, *the polynomial $y = F(x)$ furnished by the first $n$ terms of Taylor's expansion of $f(x)$ is a special interpolation polynomial, of degree not greater than $n-1$,* such that the graph of $y = F(x)$ and of $y = f(x)$ have at a given point a contact of order $n-1$.

The next important special case is when

$$\alpha_1 = \alpha_2 = \ldots = \alpha_m = 2.$$

This is known as the *Hermite interpolation formula.* Here we represent the polynomial (28) in a different form; namely as a linear combination of the given values

$$f(x_1), \quad \ldots, \quad f(x_m), \qquad f'(x_1), \quad \ldots, \quad f'(x_m).$$

We find

$$\begin{aligned} F(x) &= \sum_{i=1}^{m} h_i(x)f(x_i)+\sum_{i=1}^{m} \bar{h}_i(x)f'(x_i), \\ h_i(x) &= \left[1-\frac{\omega''(x_i)}{\omega'(x_i)}(x-x_i)\right]\{\omega_i(x)\}^2, \\ \bar{h}_i(x) &= (x-x_i)\omega_i(x), \end{aligned} \tag{32}$$

where $\omega(x)$ and $\omega_i(x)$ are given by (5) and (6).

$$f(x) = F(x)+R_{2m}(x), \qquad R_{2m}(x) = \frac{f^{(2m)}(\xi)}{(2m)!}\{\omega(x)\}^2. \tag{33}$$

The functions $h_i(x)$, $\bar{h}_i(x)$ *are called the fundamental Hermitian functions of first and second kind respectively.*

In order to prove (32) we must show that $F(x)$ is of degree less than or equal to $2m-1$ and satisfies the relations

$$F^{(s)}(x_i) = f^{(s)}(x_i), \quad s = 0, 1, \quad i = 1, 2,\ldots, m. \tag{34}$$

In fact, by (7) and (32),

$$h_i(x_j) = \begin{cases} 0, & j \neq i, \\ 1, & j = i, \end{cases} \quad i, j = 1, 2,\ldots, m,$$

$$\bar{h}_i(x_j) = 0,$$

$$h_i'(x_j) = \frac{\omega''(x_i)}{\omega'(x_i)}\{\omega_i(x_j)\}^2 + \left[1 - \frac{\omega''(x_i)}{\omega'(x_i)}(x_j - x_i)\right] 2\omega_i(x_j)\omega_i'(x_j).$$

From which

$$h_i'(x_j) = 0, \quad j \neq i,$$

$$h_i'(x_i) = -\frac{\omega''(x_i)}{\omega'(x_i)} + 2\frac{d}{dx}\omega_i(x)\Big]_{x=x_i}$$

$$= -\frac{\omega''(x_i)}{\omega'(x_i)} + \frac{2}{\omega'(x_i)}\frac{d}{dx}\frac{\omega(x)}{(x-x_i)}\Big]_{x=x_i}.$$

By Taylor's formula, since $\omega(x_i) = 0$,

$$\frac{\omega(x)}{x-x_i} = \omega'(x_i) + \frac{(x-x_i)}{2!}\omega''(x_i) + \frac{(x-x_i)^2}{3!}\omega'''(x_i) + \dots.$$

We find
$$\frac{d}{dx}\left[\frac{\omega(x)}{x-x_i}\right]_{x=x_i} = \frac{\omega''(x_i)}{2},$$

so that

$$h_i'(x_i) = -\frac{\omega''(x_i)}{\omega'(x_i)} + \frac{\omega''(x_i)}{\omega'(x_i)} = 0, \quad i = 1, 2, \dots, m.$$

Furthermore
$$\bar{h}_i'(x_j) = \begin{cases} 0, & j \neq i, \\ 1, & j = i. \end{cases}$$

The validity of (34) is thus established. Note that $R_{2m}(x) = 0$ if $f(x)$ is a polynomial of degree not greater than $2m-1$. Taking $f(x) = 1$, we derive the important identity

$$\sum_{i=1}^{m} h_i(x) \equiv 1. \tag{35}$$

Note further that in $R(x)$, as given by (33), $f^{2m}(\xi)$ is multiplied by a *non-negative factor*. Hence the remarks made above concerning the remainder in (21) are applicable here, as well as in the general case (31) if all $\alpha_i$ are even integers.

## 4. Interpolation formulae of Gauss, Stirling, and Bessel

If the independent variable $x$ takes on equidistant values it is often advantageous to introduce a new independent variable $t = (x-a)/h$, so that the set of values $x = a, a+h, \dots, a+nh$ corresponds to $t = 0, 1, 2, \dots, n$.

With such a substitution, Newton's interpolation formula (20) becomes

$$F(t) = f(a+th) = F(0)+\frac{t^{(1)}}{1!}\Delta F(0)+\frac{t^{(2)}}{2!}\Delta^2 F(0)+\ldots+$$

$$+\frac{t^{(n-1)}}{(n-1)!}\Delta^{n-1}F(0)+R_n(t), \qquad (36)$$

$$R_n(t) = \frac{t^{(n)}}{n!}F^{(n)}(\xi).$$

Here $\xi$ lies between the smallest and the largest of the numbers 0, 1,..., $n$ and $t$.

*(a) Gauss's interpolation formulae*

Various interpolation formulae can be derived from (18) by properly specifying and arranging the given points $x$. The most important are the following ones.

We introduce here the following notation: $x^{[0]} = 1$,

$$x^{[2\nu]} = x^2(x^2-1)(x^2-4)\ldots\{x^2-(\nu-1)^2\}, \quad \nu > 0,$$

$$x^{[2\nu+1]} = x\left(x^2-\frac{1}{4}\right)\left(x^2-\frac{9}{4}\right)\ldots\left(x^2-\frac{(2\nu-1)^2}{4}\right),$$

$$x^{[2\nu]-1} = \frac{x^{[2\nu]}}{x} = x(x^2-1)(x^2-4)\ldots\{x^2-(\nu-1)^2\},$$

$$x^{[2\nu+1]-1} = \frac{x^{[2\nu+1]}}{x} = \left(x^2-\frac{1}{4}\right)\left(x^2-\frac{9}{4}\right)\ldots\left(x^2-\frac{(2\nu-1)^2}{4}\right).$$

Then

$$F(t) = f(a+th) = F(0)+\frac{t}{1!}\Delta F(0)+\frac{t(t-1)}{2!}\Delta^2 F(-1)+$$

$$+\frac{(t+1)t(t-1)}{3!}\Delta^3 F(-1)+\frac{(t+1)t(t-1)(t-2)}{4!}\Delta^4 F(-2)+\ldots+$$

$$+\frac{(t+k-1)(t+k-2)\ldots t(t-1)\ldots(t-k)}{(2k)!}\Delta^{2k}F(-k)+R_{2k}$$

$$= \sum_{i=0}^{k}\left[\frac{t^{[2i]-1}}{(2i-1)!}\Delta^{2i-1}F(-i)+\frac{(t+i)t^{[2i]-1}}{(2i)!}\Delta^{2i}F(-i)\right]+R_{2k}, \quad (37)$$

$$R_{2k} = \frac{t^{[2k+2]-1}}{(2k+1)!}F^{(2k+1)}(\tau), \quad \left(\frac{1}{(-1)!} = 0\right);$$

$$F(t) = \sum_{i=0}^{k-1} \left[\frac{(t+i)t^{[2i]-1}}{(2i)!}\Delta^{2i}F(-i)+\frac{t^{[2i+2]-1}}{(2i+1)!}\Delta^{2i+1}F(-i-1)\right]+ \\ +R_{2k-1},$$

$$R_{2k-1} = \frac{(t+k)t^{[2k]-1}}{(2k)!}F^{(2k)}(\tau) \quad \left(\frac{1}{0!}=1\right).$$

This is the first Gaussian formula stopped at the order $2k$ or $2k-1$ respectively. It is derived from (18) by using the following set of $2k+1$ values:

$$x = a-kh, a-(k-1)h, ..., a-h, a, a+h, ..., a+kh \tag{38}$$

arranged in the following order:

$$a,\quad a+h,\quad a-h,\quad a+2h,\quad a-2h,\quad ...,\quad a+kh,\quad a-kh.$$

Similarly, arranging the same points (38) in the order

$$a,\quad a-h,\quad a+h,\quad a-2h,\quad a+2h,\quad ...,\quad a-kh,\quad a+kh,$$

we arrive at the second Gaussian formula, where we again stop at the order $2k$ or $2k-1$:

$$\left.\begin{aligned} F(t) &= \sum_{i=0}^{k} \left[\frac{t^{[2i]-1}}{(2i-1)!}\Delta^{2i-1}F(-i+1)+\right. \\ &\qquad\left.+\frac{(t-i)t^{[2i]-1}}{(2i)!}\Delta^{2i}F(-i)\right]+R_{2k}, \\ R_{2k} &= \frac{t^{[2k+2]-1}}{(2k+1)!}F^{(2k+1)}(\tau); \\ F(t) &= \sum_{i=0}^{k-1} \left[\frac{(t-i)t^{[2i]-1}}{(2i)!}\Delta^{2i}F(-i)+\frac{t^{[2i+2]-1}}{(2i+1)!}\Delta^{2i+1}F(-i)\right]+ \\ &\qquad +R_{2k-1}, \\ R_{2k-1} &= (t-k)\frac{t^{[2k]-1}}{(2k)!}F^{(2k)}(\tau). \end{aligned}\right\} \tag{39}$$

Similar reasoning applies if we are given an even number of values of $x$:

$$x = a-(k-1)h, a-(k-2)h, ..., a-h, a, a+h, ..., a+kh,$$

which we arrange as

$$a,\quad a+h,\quad a-h,\quad a+2h,\quad a-2h,\quad ...,\quad a-(k-1)h,\quad a+kh$$

or as

$$a+h,\quad a,\quad a+2h,\quad a-h,\quad ...,\quad a+kh,\quad a-(k-1)h.$$

We obtain, stopping at the order $2k-1$ or $2k$, four more Gaussian formulae, like (37) and (39). We write down only the formulae obtained when stopping at the order $2k-1$, where we replace $F(t)$ by $F(t+\frac{1}{2})$:

$$\begin{gathered} F(t+\tfrac{1}{2}) = \sum_{i=0}^{k-1} \left[\frac{t^{[2i+1]-1}}{(2i)!} \Delta^{2i} F(-i+1) + \right. \\ \left. + \frac{(t-i-\frac{1}{2}) t^{[2i+1]-1}}{(2i+1)!} \Delta^{2i+1} F(-i)\right] + R_{2k-1}, \\ R_{2k-1} = \frac{t^{[2k+1]-1}}{(2k)!} F^{(2k)}(\tau); \\ F(t+\tfrac{1}{2}) = \sum_{i=0}^{k-1} \left[\frac{t^{[2i+1]-1}}{(2i)!} \Delta^{2i} F(-i) + \right. \\ \left. + \frac{(t+i+\frac{1}{2}) t^{[2i+1]-1}}{(2i+1)!} \Delta^{2i+1} F(-i)\right] + R_{2k-1}, \\ R_{2k-1} = \frac{t^{[2k+1]-1}}{(2k)!} F^{(2k)}(\tau). \end{gathered} \tag{40}$$

(*b*) *Stirling's interpolation formula*

This is derived by taking the arithmetical mean of the first formulae in (37) and (39):

$$\begin{gathered} F(t) = \sum_{i=0}^{k} \left[\frac{t^{[2i]-1}}{(2i-1)!} \frac{\Delta^{2i-1} F(-i) + \Delta^{2i-1} F(-i+1)}{2} + \right. \\ \left. + \frac{t^{[2i]}}{(2i)!} \Delta^{2i} F(-i)\right] + R_{2k}, \end{gathered} \tag{41}$$

$$R_{2k} = \frac{t^{[2k+2]-1}}{(2k+1)!} F^{(2k+1)}(\tau).$$

In order to obtain the expression for $R_{2k}$ in (41) we replace $\tau$ in (37) and (39) by $\tau_1$ and $\tau_2$ respectively. Let

$$m \leqslant F^{(2k+1)}(t) \leqslant M$$

for the values of $t$ under consideration. Then

$$m \leqslant F^{(2k+1)}(\tau_i) \leqslant M \quad (i = 1, 2),$$

$$m \leqslant \frac{F^{(2k+1)}(\tau_1) + F^{(2k+1)}(\tau_2)}{2} \leqslant M.$$

It follows, $F^{(2k+1)}(t)$ being continuous, that

$$\frac{F^{(2k+1)}(\tau_1)+F^{(2k+1)}(\tau_2)}{2} = F^{(2k+1)}(\tau).$$

(c) *Bessel's interpolation formula*

This is derived by taking the arithmetic mean of the formulae (40):

$$F(t+\tfrac{1}{2}) = \sum_{i=0}^{k-1}\left[\frac{t^{[2i+1]-1}}{(2i)!}\,\frac{\Delta^{2i}F(-i)+\Delta^{2i}F(-i+1)}{2}+\right.$$

$$\left.+\frac{t^{[2i+1]}}{(2i+1)!}\Delta^{2i+1}F(-i)\right]+R_{2k-1}, \quad (42)$$

$$R_{2k-1} = \frac{t^{[2k+1]-1}}{(2k)!}F^{(2k)}(\tau).$$

## 5. Inverse interpolation

By this we mean finding the value of the independent variable $(x)$ corresponding to a given intermediate value of the function $y = f(x)$. This can be done by using the interpolation formulae with $x$ and $f(x)$ interchanged. It is sometimes preferable to use an interpolation formula, considering it as an equation in $x$, and solve it for $x$. This procedure is simple and convenient when we use linear interpolation, as was done above in finding antilogarithms.

*Illustration.* For a certain function $f(x)$ the following table is given:

| $x$ | 0 | 0·25 | 0·5 | 0·75 |
|---|---|---|---|---|
| $y$ | 0 | 0·015625 | 0·12500 | 0·421875 |

What value of $x$ gives $y = 0{\cdot}314432$ ?

Using Newton's interpolation formula (20), with $n = 2$, we get

$$y = 0{\cdot}314432 = 0+\frac{0{\cdot}015625}{0{\cdot}25}x+\frac{0{\cdot}093750x(x-0{\cdot}25)}{0{\cdot}0625}.$$

Our problem is thus reduced to finding the positive root of the quadratic equation

$$1{\cdot}5x^2-0{\cdot}3125x-0{\cdot}314432 = 0.$$

We can proceed, however, by a different method. Using the Lagrange interpolation formula inversely, we write

$$x = \frac{\omega(y)}{(y-0)\omega'(0)}0+\frac{\omega(y)}{(y-y_2)\omega'(y_2)}y_2+$$
$$+\frac{\omega(y)}{(y-y_3)\omega'(y_3)}y_3+\frac{\omega(y)}{(y-y_4)\omega'(y_4)}y_4,$$

where $$\omega(y) = y(y-y_2)(y-y_3)(y-y_4).$$

We must find $x$ for $y = 0{\cdot}314432$.

As a matter of fact, the values given are those of the function $y = x^3$, and the exact value is $x = 0{\cdot}68$. Observe that inverse interpolation may lack a simple formula for its remainder, for even if $f(x)$ is a simple function, with simple expressions for its derivatives, this may not be the case for the inverse function.

B. MECHANICAL QUADRATURES

## 6. Statement of the problem

By a mechanical quadratures formula is meant a formula which gives an approximate value for

$$\int_a^b f(x)\,dx,$$

or more generally, $$\int_a^b p(x)f(x)\,dx,$$

where $p(x)$ is a weight function. The usual result is a linear combination of the functional values $f(x_i)$ at certain preassigned points in $(a, b)$. Thus

$$\int_a^b f(x)\,dx \simeq \sum_{i=1}^{n} \lambda_i f(x_i), \tag{43}$$

where $x_i$ are the 'abscissae'

$$a \leqslant x_1 < x_2 < \ldots < x_n \leqslant b \tag{44}$$

and $\lambda_i$ are the 'coefficients', sometimes called Cotes' numbers of the mechanical quadratures formula. It is important to observe that the $x_i$ are independent of the function so that once

the abscissae are given one computation of $\lambda_i$ is sufficient for all functions.

We write

$$\int_a^b p(x)f(x)\,dx = \sum_{i=1}^{n} \lambda_i f(x_i) + R_n(f)$$
$$\equiv Q_n(f) + R_n(f).$$

The remainder $R_n(f)$ may vanish for certain functions. In the cases that we are to discuss $R_n(f) = 0$ when $f(x)$ is a polynomial of not too great degree.

It is possible to construct a more complicated mechanical quadratures formula involving linear aggregates of various derivatives. However, in practice such a formula when $f(x)$ is given in graphical or tabular form is of very limited value. For this reason we confine our discussion to formulae of the type (43) except for a brief discussion of a form which arises from Hermite's interpolation formula (32), involving $f(x_i)$ and $f'(x_i)$.

Observe further that the abscissae may be taken from a given infinite set of points, e.g. from an infinite array

$$\begin{matrix} x_{11} & & & & & & & \\ x_{12} & x_{22} & & & & & & \\ x_{13} & x_{23} & x_{33} & & & & & \\ . & . & . & . & . & . & . & \\ x_{1n} & x_{2n} & . & . & . & . & x_{nn} & \\ . & . & . & . & . & . & . & . \end{matrix} \tag{45}$$

in which case we obtain correspondingly an infinite sequence of mechanical quadratures formulae

$$\int_a^b p(x)f(x)\,dx = \sum_{i=1}^{n} \lambda_{in} f(x_{in}) + \rho_n(f). \tag{46}$$

The question of *convergence* then arises, namely: *Given the infinite sequence of abscissae $x_{in}$, $i = 1, 2,..., n$; $n = 1, 2,...$; for what $f(x)$ is it true that*

$$\lim_{n\to\infty} \rho_n(f) = 0\,?$$

This will be but briefly discussed in what follows.

## 7. Mechanical quadratures and interpolation

Interpolation is the most natural source of mechanical quadratures formulae.

First, consider the Lagrange interpolation formula (12). By integration, we obtain from it the fundamental mechanical quadratures formula

$$\int_a^b f(x)\,dx = \int_a^b L_{n-1}(x)\,dx + \rho_n(f)$$

$$= \sum_{i=1}^n H_i f(x_i) + \rho_n(f), \tag{47}$$

$$H_i = \int_a^b \frac{\omega(x)\,dx}{(x-x_i)\omega'(x_i)} \equiv \int_a^b \omega_i(x)\,dx, \tag{48}$$

$$\rho_n(f) = \int_a^b R_n(x)\,dx = \frac{1}{n!}\int_a^b \omega(x) f^{(n)}(\xi)\,dx. \tag{49}$$

This integral in (49) exists by (47) inasmuch as $f(x)$ and $L_n(x)$ are integrable. More generally, introducing a weight-factor $p(x)$,

$$\int_a^b p(x)f(x)\,dx = \int_a^b p(x)L_{n-1}(x)\,dx = \sum_{i=1}^n \bar{H}_i f(x_i) + \bar{\rho}_n(f), \tag{50}$$

$$\bar{H}_i = \int_a^b p(x)\omega_i(x)\,dx, \tag{51}$$

$$\bar{\rho}_n(f) = \int_a^b p(x)R_n(x)\,dx = \frac{1}{n!}\int_a^b p(x)\omega(x)f^{(n)}(\xi)\,dx. \tag{52}$$

Similarly from the general interpolation formula (31), where $F(x)$ and $R(x)$ are given by (28) and (31) respectively, we generate a more complicated mechanical quadratures formula

$$\int_a^b f(x)\,dx = \int_a^b F(x)\,dx + \rho_n(f), \tag{53}$$

$$\rho_n(f) = \int_a^b R_n(x)\,dx = \frac{1}{n!}\int_a^b (x-x_1)^{\alpha_1}\dots(x-x_n)^{\alpha_m} f^{(n)}(\xi)\,dx.$$

In particular,

$$\int_a^b f(x)\,dx = \sum_{i=1}^{n} H_i f(x_i) + \sum_{i=1}^{n} \bar{H}_i f'(x_i) + \rho_n(f). \tag{54}$$

This is derived from Hermite's interpolation formula (32). Here

$$\begin{aligned} H_i &= \int_a^b h_i(x)\,dx = \int_a^b \left[1 - \frac{\omega''(x_i)}{\omega'(x_i)}\right]\{\omega_i(x)\}^2\,dx, \\ \bar{H}_i &= \int_a^b \bar{h}_i(x)\,dx = \int_a^b (x-x_i)\{\omega_i(x)\}^2\,dx \\ &= \int_a^b \omega(x)\frac{\omega_i(x)}{\omega'(x_i)}\,dx, \quad i = 1, 2,..., n, \end{aligned} \tag{55}$$

$$\rho_n(f) = \frac{1}{(2n)!}\int_a^b \{\omega(x)\}^2 f^{(2n)}(\xi)\,dx = \frac{f^{(2n)}(\eta)}{(2n)!}\int_a^b \{\omega(x)\}^2\,dx. \tag{56}$$

Here we integrate the remainder formula and then apply the first law of the mean. We have remarked that the remainder is continuous in $x$. Hence $f^{(2n)}(\xi)$ is continuous in $x$ where $x \neq x_1, x_2,..., x_n$. However, it approaches a limit when $x$ approaches each of these points and consequently can be defined at them so as to render $f^{(2n)}(\xi)$ continuous in $x$ over the entire interval $(a, b)$, hence the application of the mean-value theorem is justified.

Here $\eta$, like $\xi$, denotes a certain number lying in $(a, b)$ and depending on $n$ and $f(x)$.

In (54), (55), (56) the integrands involved may have a weight-factor $p(x)$.

## 8. Degree of precision

*A mechanical quadratures formula which is exact for an arbitrary polynomial, $G_k(x)$, of degree less than or equal to $k$ is said to have a 'degree of precision' equal to $k$.*

Since the Lagrange interpolation formula is exact for any

polynomial of degree $n-1$, we conclude that *the degree of precision of a mechanical quadratures formula generated by the Lagrange interpolation formula, with $n$ arbitrarily preassigned abscissae, is at least $n-1$.* Moreover, *it cannot exceed $2n-1$.* For, let the given polynomial be $\{\omega(x)\}^2$, then

$$\int_a^b \{\omega(x)\}^2\, dx = \sum_{i=1}^n H_i\{\omega(x_i)\}^2.$$

But $\omega(x_i) = 0$. Hence

$$\int_a^b \{\omega(x)\}^2\, dx = 0.$$

This is impossible since $\omega(x) \not\equiv 0$.

Formula (54) has a degree of precision equal to $2n-1$, as is seen from the expression (56) for its remainder. We shall see below the interesting fact that the maximum degree of precision, $2n-1$, is actually attained by an extensive class of mechanical quadratures formulae of the simplest type (47), provided the abscissae are properly chosen.

## 9. Cotes' formula. Rule of rectangles. Trapezoidal rule. Simpson's rule. Tchebycheff's formula.

By specifying the number and the location of the abscissae in (47) various specialized formulae are derived. We proceed to discuss the most important ones.

Observe that if the interval of integration $(a, b)$ is finite, it can always be reduced to $(0, 1)$ or to $(-1, 1)$ by means of the respective substitutions

$$x = a+(b-a)y, \qquad x = a+\frac{b-a}{2}(y+1), \tag{57}$$

resulting in

$$\int_a^b f(x)\, dx = (b-a)\int_0^1 f[a+(b-a)y]\, dy$$

$$\text{or} \qquad \int_a^b f(x)\, dx = \frac{b-a}{2}\int_{-1}^1 f\left[a+\frac{b-a}{2}(y+1)\right] dy, \tag{58}$$

(i) *Cotes' formula*

Cotes' formula corresponds to *equidistant abscissae.* Taking for the interval of integration (0, 1), we have

$$x_1 = 0, \quad x_2 = \frac{1}{n-1}, \quad x_3 = \frac{2}{n-1}, \quad \dots, \quad x_n = \frac{n-1}{n-1} = 1,$$

$$\int_0^1 f(x)\,dx \simeq \sum_{i=1}^{n} H_{in} f\left(\frac{i-1}{n-1}\right), \tag{59}$$

$$H_{in} = (n-1)^{n-1} \int_0^1 \frac{x\left(x-\frac{1}{n-1}\right)\dots\left(x-\frac{i-1}{n-1}\right)\left(x-\frac{i+1}{n-1}\right)\dots(x-1)\,dx}{(i-1)(i-2)\dots2.1(-1)(-2)\dots(i-n)}$$

$$= \frac{(-1)^{n-i}}{n-1} \int_0^{n-1} \frac{t(t-1)\dots(t-i+1)(t-i-1)\dots(t-n+1)\,dt}{\{(i-1)!\}\{(n-i)!\}}.$$

A mechanical quadratures formula might be called 'simple' if either its abscissae or coefficients are given according to a simple law. Thus, Cotes' formula is 'simple'. However, gaining in simplicity with regard to abscissae we lost in other respects. First, the form of the remainder for (59) as given below, namely,

$$\rho_n(f) = \frac{1}{n!} \int_a^b (x-a)\left(x-a-\frac{b-a}{n-1}\right)\dots(x-b) f^{(n)}(\xi)\,dx, \tag{60}$$

generally cannot be improved and admits of a rough estimate only:

$$|\rho_n(f)| \leqslant \frac{M_n}{n!} \int_a^b \left|(x-a)\left(x-a-\frac{b-a}{n-1}\right)\dots(x-b)\right| dx,$$
$$|f^{(n)}(x)| \leqslant M_n, \quad a \leqslant x \leqslant b. \tag{61}$$

The second disadvantage lies in the behaviour of (59) as $n \to \infty$. Integration is a 'smoothing out process', and consequently one might expect that even if the original interpolation formula diverges, it might generate a convergent formula of mechanical

quadratures. This expectation is justified in many cases, but *not in the case of Cotes' formula.*

(ii) *Rule of rectangles*

In (47) let $n = 1$ and $x_i = a,\ b,\ \frac{a+b}{2}$ successively. We find the following formulae:

$$\int_a^b f(x)\,dx = (b-a)f(a)+\frac{(b-a)^2}{2}f'(\eta_1), \tag{62}$$

$$\int_a^b f(x)\,dx = (b-a)f(b)-\frac{(b-a)^2}{2}f'(\eta_2), \tag{63}$$

$$\int_a^b f(x)\,dx = (b-a)f\left(\frac{a+b}{2}\right)+\rho(f). \tag{64}$$

These are frequently called the rules of interior rectangles, exterior rectangles, and middle rectangles, respectively.

The formulae for the remainders in (62) and (63) are derived by applying the first theorem of the mean to

$$\int_a^b (x-a)f'(\xi_1)\,dx \quad \text{and} \quad \int_a^b (x-b)f'(\xi_2)\,dx$$

since $x-a$ and $x-b$ each keeps a constant sign in $(a, b)$. The reasoning here is the same as on p. 96. However, it is not possible to apply the first theorem of the mean in (64) where the remainder is

$$\int_a^b \left(x-\frac{a+b}{2}\right)f'(\xi_3)\,dx.$$

Here $\left(x-\frac{a+b}{2}\right)$ does not have a fixed sign when $a < x < b$. In order to derive an improved formula for the remainder in this case, we proceed as follows. By Taylor's formula,

$$f(x) = f\left(\frac{a+b}{2}\right)+\left(x-\frac{a+b}{2}\right)f'\left(\frac{a+b}{2}\right)+\frac{\left(x-\frac{a+b}{2}\right)^2}{2!}f''(\xi);$$

whence

$$\int_a^b f(x)\,dx = (b-a)f\left(\frac{a+b}{2}\right)+f'\left(\frac{a+b}{2}\right)\int_a^b \left(x-\frac{a+b}{2}\right)dx+\rho(f),$$

$$\rho(f) = \frac{1}{2}\int_a^b \left(x-\frac{a+b}{2}\right)^2 f''(\xi)\,dx.$$

Observing that

$$\int_a^b \left(x-\frac{a+b}{2}\right)dx = 0, \qquad \int_a^b \left(x-\frac{a+b}{2}\right)^2 dx = \frac{(b-a)^3}{12},$$

we get (64) with its remainder in the following form:

$$\int_a^b f(x)\,dx = (b-a)f\left(\frac{a+b}{2}\right)+\frac{(b-a)^3}{24}f''(\eta). \tag{65}$$

The term 'rules of rectangles' is justified by the fact that if $f(x)$ is non-negative for $a < x < b$, and if the graph of $y = f(x)$ is plotted in the ordinary cartesian plane, then our formulae are equivalent to replacing the area between the curve $y = f(x)$ and the $x$-axis from $x = a$ to $x = b$ by the area of a rectangle.

(iii) *Trapezoidal rule*

Here we replace the area under the curve by the area of a trapezoid.

The corresponding mechanical quadratures formula is derived by letting $n = 2$, that is, $x_1 = a$, $x_2 = b$ in Cotes' formula. We thus get, again applying the mean-value theorem to

$$\int_a^b (x-a)(x-b)f''(\xi)\,dx,$$

$$\int_a^b f(x)\,dx = \frac{b-a}{2}[f(a)+f(b)]-\frac{(b-a)^3}{12}f''(\eta). \tag{66}$$

The right to apply the mean-value theorem in a like case was discussed on p. 96.

It is seen from (65) and (66) that if $f''(x)$ keeps a constant sign in $(a, b)$ then $\int\limits_a^b f(x)\,dx$ lies between

$$\frac{b-a}{2}[f(a)+f(b)] \quad \text{and} \quad \frac{b-a}{2}f\left(\frac{a+b}{2}\right).$$

(iv) *Simpson's rule*

In Cotes' formula let

$$n = 3, \quad x_1 = a, \quad x_2 = \frac{a+b}{2}, \quad x_3 = b.$$

Here the area under the curve is replaced by that under a parabola. We get

$$\int\limits_a^b f(x)\,dx = \frac{b-a}{6}\left[f(a)+4f\left(\frac{a+b}{2}\right)+f(b)\right]+\rho(f). \tag{67}$$

The remainder, as derived from (49), is given by

$$\rho(f) = \frac{1}{6}\int\limits_a^b (x-a)\left(x-\frac{a+b}{2}\right)(x-b)f'''(\xi)\,dx.$$

This is not a convenient form. We consequently try the following special case of formula (53):

$$x_1 = a, \quad \alpha_1 = 1, \quad x_2 = \frac{a+b}{2}, \quad \alpha_2 = 2, \quad x_3 = b, \quad \alpha_3 = 1;$$

$$n = 1+2+1 = 4.$$

This proves to give the same interpolation polynomial but with a different formula for the remainder. We find

$$\rho(f) = \frac{1}{4!}\int\limits_a^b (x-a)\left(x-\frac{a+b}{2}\right)^2(x-b)f^{(\mathrm{iv})}(\xi)\,dx$$

$$= \frac{1}{4!}f^{(\mathrm{iv})}(\eta)\left(\frac{b-a}{2}\right)^5\int\limits_{-1}^{1} t^2(t^2-1)\,dt = -\left(\frac{b-a}{2}\right)^5\frac{f^{(\mathrm{iv})}(\eta)}{90}$$

since we can apply the mean-value theorem. We obtain

$$\int\limits_a^b f(x)\,dx = \frac{b-a}{6}\left[f(a)+4f\left(\frac{a+b}{2}\right)+f(b)\right]\frac{1}{90}\left(\frac{b-a}{2}\right)^5 f^{(\mathrm{iv})}(\eta). \tag{68}$$

This is 'Simpson's rule' much used in practical computations.

We conclude from (67) that its degree of precision is 3. This is because $f^{(iv)}(x) \equiv 0$ if $f(x)$ is a polynomial of degree 3.

The error arising from the use of the formulae given under (i), (ii), (iii), (iv) can usually be lowered by subdividing the given interval $(a, b)$ into a number, $s$, of sub-intervals of equal length, so that

$$\int\limits_a^b f(x)\,dx = \int\limits_a^{a+(b-a)/s} f(x)\,dx + \int\limits_{a+(b-a)/s}^{a+2(b-a)/s} f(x)\,dx + \ldots + \int\limits_{b-(b-a)/s}^{b} f(x)\,dx$$

and applying to each integral on the right the mechanical quadratures formulae with their remainders.

To illustrate, we use formula (62). Let

$$f_0 = f(a),\quad f_1 = f\left(a+\frac{b-a}{s}\right),\quad \ldots,\quad f_i = f\left(a+i\frac{b-a}{s}\right),\quad \ldots,$$
$$f_s = f(b).$$

We get

$$\int\limits_a^b f(x)\,dx = \frac{b-a}{s}[f_0+f_1+\ldots+f_{s-1}]+\rho(f), \tag{69}$$

$$\rho(f) = \frac{(b-a)^2}{2s^2}[f'(\eta_1)+f'(\eta_2)+\ldots+f'(\eta_s)],$$

$$a < \eta_1 < a+\frac{b-a}{s} < \eta_2 < \ldots < \eta_s < b.$$

Moreover, if $\quad m \leqslant f'(x) \leqslant M, \quad a \leqslant x \leqslant b,$

then $\quad sm \leqslant f'(\eta_1)+\ldots+f'(\eta_s) \leqslant sM,$

whence $\quad \dfrac{f'(\eta_1)+\ldots+f'(\eta_s)}{s} = f'(\eta)$

and (69) gives

$$\int\limits_a^b f(x)\,dx = \frac{b-a}{s}[f_0+f_1+\ldots+f_{s-1}]+\frac{(b-a)^2}{2s}f'(\eta).$$

**If we compare this with the remainder form in (62) we immediately notice the $s$ in the denominator and consequently expect a smaller remainder.**

Similarly,

$$\int_a^b f(x)\,dx = \frac{b-a}{s}[f_1+f_2+...+f_s]-\frac{(b-a)^2}{2s}f'(\eta), \tag{70}$$

$$\int_a^b f(x)\,dx = \frac{b-a}{s}[f_{\frac{1}{2}}+f_{\frac{3}{2}}+...+f_{s-\frac{1}{2}}]+\frac{(b-a)^3}{24s^2}f''(\eta), \tag{71}$$

$$\left[f_{\frac{1}{2}i} = f\left\{a+\frac{i(b-a)}{2s}\right\}\right],$$

$$\int_a^b f(x)\,dx = \frac{b-a}{2s}[f_0+2f_1+2f_2+...+2f_{s-1}+f_s]-\frac{(b-a)^3}{12s^2}f''(\eta), \tag{72}$$

$$\int_a^b f(x)\,dx = \frac{b-a}{6s}[f_0+4f_{\frac{1}{2}}+2f_1+4f_{\frac{3}{2}}+2f_2+...+$$
$$+2f_{s-1}+4f_{s-\frac{1}{2}}+f_s]-\frac{1}{90s^4}\left(\frac{b-a}{2}\right)^5 f^{(\mathrm{iv})}(\eta). \tag{73}$$

Note the presence of the factor $1/s^4$ in the last remainder.

Some of the above formulae involve the functional values at the end-points of the interval of integration, some use interior points only. Accordingly they are often classified as mechanical quadratures formulae of the *closed* or *open* type.

(v) *Tchebycheff's formula*

This is another 'simple' formula of mechanical quadratures, where *the abscissae are so chosen that all coefficients are equal.* Thus,

$$\int_a^b f(x)\,dx = H[f(x_1)+f(x_2)+...+f(x_n)], \tag{74}$$

which makes it especially useful in applications. Having $n+1$ constants at our disposal, $x_1, x_2,..., x_n$ and $H$, we hope to build (74) *with degree of precision at least n.* For this the following $n+1$ equations must be satisfied:

$$\int_a^b x^0\,dx = b-a = nH,$$

$$\int_a^b x^1\,dx = H(x_1+x_2+...+x_n), \tag{75}$$

$$\cdots\cdots\cdots\cdots$$

$$\int_a^b x^n\,dx = H(x_1^n+x_2^n+...+x_n^n).$$

If $(a, b)$ has been reduced to $(-1, 1)$, we get

$$\int_{-1}^{1} f(x)\,dx = \frac{2}{n}[f(x_1)+f(x_2)+\ldots+f(x_n)], \tag{76}$$

where the abscissae $x_i$ are to be determined from the following system of equations:

$$\begin{aligned} x_1+x_2+\ldots+x_n &= 0,\\ x_1^2+x_2^2+\ldots+x_n^2 &= \tfrac{1}{3}n,\\ x_1^3+x_2^3+\ldots+x_n^3 &= 0,\\ \cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot &\quad\cdot\\ x_1^n+x_2^n+\ldots+x_n^n &= \begin{cases} \dfrac{n}{n+1}, & n \text{ even},\\ 0, & n \text{ odd}.\end{cases} \end{aligned} \tag{77}$$

The $x_i$ thus determined must be real, distinct, and lie in $(-1, 1)$.

Will the system (77) yield $x_i$ of the desired character, for any $n$? S. Bernstein proved that for $n > 9$, (77) yields imaginary values for some $x_i$. This limits the usefulness of Tchebycheff's formula. Another disadvantage is its unworkable remainder.

## 10. Gaussian mechanical quadratures formulae

Gauss conceived the idea of *choosing the abscissae in* (47) *so as to attain the maximum degree of precision*, $2n-1$, that is, such that

$$\int_a^b G_{2n-1}(x)\,dx = \sum_{i=1}^{n} H_i\, G_{2n-1}(x_i) \tag{78}$$

whenever the degree of $G_{2n-1}(x)$ is less than or equal to $2n-1$. Generalizing Gauss's reasoning, we proceed to show how the same end can be achieved in (50), that is, we solve the following problem:

Choose the abscissae $x_1$,

$$a \leqslant x_1 < x_2 < \ldots < x_n \leqslant b,$$

*so that* (50) *shall have a degree of precision* $2n-1$:

$$\int_a^b p(x)G_{2n-1}(x)\,dx = \sum_{i=1}^{n} H_i\, G_{2n-1}(x_i), \tag{79}$$

$$H_i = \int_a^b p(x)\frac{\omega(x)\,dx}{(x-x_i)\omega'(x_i)}, \quad \omega = (x-x_1)\ldots(x-x_n),$$

where $p(x)$ is defined over $(a, b)$.

We prove the following theorem:

THEOREM. *A necessary and sufficient condition for the validity of* (79) *is the 'orthogonality relation'*

$$\int_a^b p(x)\omega(x)G_{n-1}(x)\,dx = 0, \tag{80}$$

*where $G_{n-1}(x)$ is any polynomial of degree $n-1$.*

Write $$G_{2n-1}(x) = \omega(x)G_{n-1}(x)+\bar{G}_{n-1}(x).$$

Hence, since $\omega(x_i) = 0$,

$$G_{2n-1}(x_i) = \bar{G}_{n-1}(x_i), \quad i = 1, 2,..., n. \tag{81}$$

$$\int_a^b p(x)G_{2n-1}(x)\,dx = \int_a^b p(x)\omega(x)G_{n-1}(x)\,dx + \int_a^b p(x)\bar{G}_{n-1}(x)\,dx.$$

If (79) holds, then

$$\sum_{i=1}^n H_i\,G_{2n-1}(x_i) = \int_a^b p(x)\omega(x)G_{n-1}(x)\,dx + \int_a^b p(x)\bar{G}_{n-1}(x)\,dx$$

$$= \int_a^b p(x)\omega(x)G_{n-1}(x)\,dx + \sum_{i=1}^n H_i\,\bar{G}_{n-1}(x_i), \tag{82}$$

for (47) is exact for a polynomial of degree $n-1$. From (82), combined with (81), the relation (80) follows directly, which proves its necessity. The proof of its sufficiency is left to the reader.

The question now arises: *Given $p(x)$, such that $\int_a^b p(x)x^n\,dx$ exists, $n = 0, 1, 2,...$, does there exist, for any $n$, a polynomial $\omega(x) = (x-x_1)...(x-x_n)$ satisfying* (80) *whose zeros are real, distinct, and in $(a, b)$?*

*The answer is in the affirmative, if $p(x)$ is non-negative in $(a, b)$ whether $(a, b)$ is finite or infinite in length. Moreover, such $\omega(x)$ is unique.*

Proof of this general theorem is omitted.

It follows that, *with such $p(x)$, we can construct mechanical quadratures formulae* (50) *with degree of precision $2n-1$, for $n = 1, 2,....$ We call these Gaussian mechanical quadratures formulae.*

We discuss here two important special cases, where the finite interval of integration has been reduced to $(-1, 1)$.

(i) $p(x) = 1$: *The Gauss case.* Here (80) becomes

$$\int_{-1}^{1} \omega(x)G_{n-1}(x)\,dx = 0. \tag{83}$$

We can consider $\omega(x)$ as the $n$th derivative of another polynomial, say $P(x)$, of degree $2n$:

$$\omega(x) = P^{(n)}(x),$$

so that (83) can be rewritten as

$$\int_{-1}^{1} P^{(n)}(x)G_{n-1}(x)\,dx = 0.$$

Repeated integration by parts gives

$$P^{(n-1)}(x)G_{n-1}(x)]_{-1}^{1} - P^{(n-2)}(x)G'_{n-1}(x)]_{-1}^{1} \pm \dots \pm P(x)G_{n-1}^{(n-1)}(x)]_{-1}^{1} \mp$$
$$\mp \int_{-1}^{1} P(x)G_{n-1}^{(n)}(x)\,dx = 0.$$

This relation, where the integral on the left vanishes, implies, in view of the arbitrariness of $G_{n-1}(x)$,

$$P^{(n-1)}(\pm 1) = P^{(n-2)}(\pm 1) = \dots = P'(\pm 1) = P(\pm 1) = 0.$$

Hence $P(x) = c(x^2-1)^n$. Choose $c$ so that

$$\omega(x) = \frac{1}{2n(2n-1)\dots(n+1)} \frac{d^n(x^2-1)^n}{dx^n} = \frac{n!}{(2n)!}\frac{d^n(x^2-1)^n}{dx^n}. \tag{84}$$

It is seen† that $P(x)$ has $n$ zeros at $x = -1$ and $n$ zeros at $x = +1$. Hence, $P'(x)$ has $n-1$ zeros at $x = -1$ and $n-i$ zeros at $x = +1$, and in addition, by Rolle's theorem, one zero between $-1$ and $+1$. Apply the same reasoning to $P''(x)$, $P'''(x)$,.... In this way we learn that $\omega(x)$ *has all zeros real, distinct, and between* $-1$ and $+1$, as required.

† Rodrigues' formula for the Legendre polynomial of order $n$ is

$$\bar{P}_n(x) = \frac{1}{(2^n)(n!)}\frac{d^n(x^2-1)^n}{dx^n} = \frac{2^n(n!)^2}{(2n)!}\omega(x).$$

We have, from (84),

$$\omega(x) = x^n - \frac{n(n-1)}{2(2n-1)}x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2.4(2n-1)(2n-3)}x^{n-4} - \dots. \tag{85}$$

The zeros, $x_i$, are symmetrical with respect to the origin,

$$x_i = -x_{n+1-i}, \quad i = 1, 2, \dots, n. \tag{86}$$

For $n$ odd, $n = 2m+1$, the middle zero $x_{m+1} = 0$.

The coefficients $H_i$, can be computed from the general formula (48). One readily verifies, by virtue of (86), that

$$H_i = H_{n+1-i}, \quad i = 1, 2, \dots, n. \tag{87}$$

Another important property of the $H_i$ is that *they are all positive*. In fact, apply (79) to the polynomial $Q_i(x) = \left[\frac{\omega(x)}{(x-x_i)\omega'(x_i)}\right]^2$ of degree $2n-2$, with

$$Q_i(x_i) = 1, \qquad Q_i(x_j) = 0, \quad j = 1, 2, \dots, i-1, i+1, \dots, n,$$

and we get

$$H_i = \int_{-1}^{1} \left[\frac{\omega(x)}{(x-x_i)\omega'(x_i)}\right]^2 dx > 0, \quad i = 1, 2, \dots, n. \tag{88}$$

Similar considerations show that *all coefficients in any Gaussian mechanical quadratures formula are positive*. This property is of decisive importance in the study of convergence of such formulae.

(ii) $p(x) = 1/\sqrt{(1-x^2)}$: *The Tchebycheff-Gauss case.* Letting $x = \cos\phi$ in (80), we get

$$\int_0^\pi \omega(\cos\phi)G_{n-1}(\cos\phi)\, d\phi = 0. \tag{89}$$

Here we take successively

$$G_{n-1}(\cos\phi) = 1, \cos\phi, \cos 2\phi, \dots, \cos(n-1)\phi,$$

since $\cos k\phi$ is a polynomial of degree $k$ in $\cos\phi$. We thus replace (89) by a system of equations as follows:

$$\int_0^\pi \omega(\cos\phi)\cos k\phi\, d\phi = 0, \quad k = 0, 1, \dots, n-1. \tag{90}$$

Now (90) will be satisfied if $\omega(\cos\phi) = C\cos n\phi$, where $C$ is a constant, because

$$\int_0^\pi \cos n\phi \cos k\phi \, d\phi = 0, \quad k \neq n. \tag{91}$$

We let

$$\omega(\cos\phi) = \frac{1}{2^{n-1}}\cos n\phi. \tag{92}$$

The zeros of $\cos n\phi$ are $\cos\phi_i = x_i$. The abscissae $x_i$ are found by letting $\cos n\phi = 0$. They are as follows:

$$x_i = -x_{n+1-i} = \cos\frac{(2i-1)\pi}{2n} \equiv \cos\phi_i; \qquad \phi_i = \frac{(2i-1)\pi}{2n}, \tag{93}$$

$$i = 1, 2, \dots, n.$$

The coefficients are equal:

$$H_1 = H_2 = \dots = H_n. \tag{94}$$

We shall prove

$$H_i = \pi/n, \quad i = 1, 2, \dots, n. \tag{95}$$

To prove (95), apply (51) and make use of (91) and (92). We get

$$2^{n-1}\omega'(x) = \frac{n\sin n\phi}{\sin\phi}, \quad x = \cos\phi,$$

$$\omega'(x_i) = \frac{n}{2^{n-1}}\frac{\sin n\phi_i}{\sin\phi_i}, \quad i = 1, 2, \dots, n. \tag{96}$$

Furthermore,

$$\frac{2^{n-1}\omega(x)}{x-x_i} = \frac{\cos n\phi}{\cos\phi-\cos\phi_i} = \frac{\cos n\phi-\cos n\phi_i}{\cos\phi-\cos\phi_i}$$

$$= \beta_0+\beta_1\cos\phi+\beta_2\cos 2\phi+\dots+\beta_{n-1}\cos(n-1)\phi, \tag{97}$$

where the coefficients $\beta_i$ are independent of $\phi$, so that, by (90) and (94),

$$H_i = \int_{-1}^{1} \frac{\omega(x)}{(x-x_i)\omega'(x_i)}\frac{dx}{\sqrt{(1-x^2)}} = \pi\beta_0\frac{\sin\phi_i}{n\sin n\phi_i}, \quad i = 1, 2, \dots, n. \tag{98}$$

It remains to determine $\beta_0$. In (97) let

$$\phi = \phi_1, \quad \phi_2 = \phi_1+\pi/n, \quad \dots, \quad \phi_n = \phi_1+(n-1)\pi/n.$$

This gives

$$\beta_0+\beta_1\cos\phi_1+...+\beta_{n-1}\cos(n-1)\phi_1 = 0,$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\beta_0+\beta_1\cos\phi_{i-1}+...+\beta_{n-1}\cos(n-1)\phi_{i-1} = 0,$$

$$\beta_0+\beta_1\cos\phi_i+...+\beta_{n-1}\cos(n-1)\phi_i = \frac{n\sin n\phi_i}{\sin\phi_i},$$

$$\beta_0+\beta_1\cos\phi_{i+1}+...+\beta_{n-1}\cos(n-1)\phi_{i+1} = 0,$$

$$\cdots\cdots\cdots\cdots\cdots$$

$$\beta_0+\beta_1\cos\phi_n+...+\beta_{n-1}\cos(n-1)\phi_n = 0.$$

Add by columns and observe that

$$\cos k\phi_1+\cos k\phi_2+...+\cos k\phi_n = 0, \quad k = 1, 2,..., n-1.$$

It results that $\beta_0 = \dfrac{\sin n\phi_i}{\sin\phi_i}.$

Substituting this expression in (98) gives the result as stated above, namely,

$$H_i = \pi/n, \quad i = 1, 2,..., n.$$

We now can show very readily that the *Tchebycheff-Gaussian formula of mechanical quadratures*

$$\int\limits_{-1}^{1} f(x)\frac{dx}{\sqrt{(1-x^2)}} = \frac{\pi}{n}\sum_{i=1}^{n} f\left(\cos\frac{(2i-1)\pi}{2n}\right)+\rho_n(f), \tag{99}$$

*converges, that is,* $\lim_{n\to\infty}\rho_n(f) = 0$, *for any bounded* $f(x)$ *for which the left-hand integral exists.* In fact, we see at once that the sum on the right side of (99) is the Riemann approximate sum for the definite integral

$$\int\limits_{0}^{\pi} f(\cos\phi)\,d\phi = \int\limits_{-1}^{1}\frac{f(x)\,dx}{\sqrt{(1-x^2)}}.$$

## 11. Remainder in Gaussian formula of mechanical quadratures

We proceed to establish the important formula

$$\int\limits_{a}^{b} p(x)f(x)\,dx = \sum_{i=1}^{n} H_i f(x_i)+\frac{f^{(2n)}(\eta)}{(2n)!}\int\limits_{a}^{b} p(x)\{\omega(x)\}^2\,dx, \tag{100}$$

*when* $\displaystyle\int\limits_{a}^{b} p(x)\omega(x)G_{n-1}(x)\,dx = 0.$

*First proof*

We know that

$$\int_a^b p(x)G_{2n-1}(x)\,dx = \sum_{i=1}^n H_i\,G_{2n-1}(x_i).$$

Subtracting this from

$$\int_a^b p(x)f(x)\,dx = \sum_{i=1}^n H_i f(x_i) + \rho_n(f) \tag{101}$$

yields

$$\rho_n(f) = \int_a^b p(x)[f(x)-G_{2n-1}(x)]\,dx - \sum_{i=1}^n H_i[f(x_i)-G_{2n-1}(x_i)]. \tag{102}$$

Substitute here for $G_{2n-1}(x)$ the Hermitian polynomial $F(x)$ such that at the points $x_i$ given in (100)

$$F(x_i) = f(x_i), \qquad F'(x_i) = f'(x_i), \quad i = 1, 2,..., n. \tag{103}$$

By (33)
$$f(x) = F(x) + \frac{f^{(2n)}(\xi)}{(2n)!}\{\omega(x)\}^2,$$

which, substituted in (102), gives (100).

*Second proof*

This constitutes a new approach to Gaussian mechanical quadratures formulae.

Consider the mechanical quadratures formula (54), where we introduce a weight-factor $p(x)$. We ask: *under what conditions is it independent of $f'(x_i)$ $(i = 1, 2,..., n)$?* The answer is furnished by the following theorem.

THEOREM. *A mechanical quadratures formula for*

$$\int_a^b p(x)f(x)\,dx,$$

*based on the Hermitian interpolation formula with n abscissae $x_i$, is independent of $f'(x_1), f'(x_2),..., f'(x_n)$ if and only if*

$$\omega(x) \equiv (x-x_1)...(x-x_n)$$

*satisfies the orthogonality relation*

$$\int_a^b p(x)\omega(x)G_{n-1}(x)\,dx = 0.$$

The sufficiency of this condition follows from the expression for $H_i$, the coefficient of $f'(x_i)$, as given in (55).

The condition is also necessary. First, it implies

$$\int_a^b p(x)\omega(x)\omega_i(x)\,dx = 0, \quad i = 1, 2,..., n. \tag{104}$$

Now, writing $\quad G_{n-1}(x) = \sum_{i=1}^{n} \omega_i(x)G_{n-1}(x_i),$

we get

$$\int_a^b p(x)\omega(x)G_{n-1}(x)\,dx = \sum_{i=1}^{n} G_{n-1}(x_i)\int_a^b p(x)\omega(x)\omega_i(x)\,dx,$$

so (104), in turn, implies (100).

Thus *any Gaussian mechanical quadratures formula is a special case of* formula (54) with a weight factor. Its remainder is therefore given by (56) with a weight-factor, or (100).

In the Gauss and Tchebycheff-Gauss cases discussed in the preceding section we can express the remainder in a more explicit form.

In Gauss's case, by (84),

$$\int_{-1}^{1} \{\omega(x)\}^2\,dx = \left[\frac{n!}{(2n)!}\right]^2 \int_{-1}^{1} \frac{d^n(x^2-1)^n}{dx^n}\,\frac{d^n(x^2-1)^n}{dx^n}\,dx.$$

Repeated integration by parts yields

$$\int_{-1} \{\omega(x)\}^2\,dx = \left[\frac{n!}{(2n)!}\right]^2 \int_{-1}^{1} (1-x^2)^n\,\frac{d^{2n}(x^2-1)^n}{dx^{2n}}\,dx$$

$$= 2\frac{(n!)^2}{(2n)!}\int_0^{\frac{1}{2}\pi} (\sin\phi)^{2n+1}\,d\phi = \frac{2(n!)^2}{(2n)!}\,\frac{2.4.6...2n}{1.3.5...(2n+1)}$$

$$= \left[\frac{n!}{1.3.5...(2n-1)}\right]^2 \frac{2}{(2n+1)!}.$$

In Tchebycheff's case, by (92),

$$\int_{-1}^{1} \frac{\{\omega(x)\}^2\,dx}{\sqrt{(1-x^2)}} = \frac{1}{2^{2(n-1)}}\int_0^{\pi} \cos^2 n\phi\,d\phi = \frac{\pi}{2^{2n-1}}.$$

We have thus derived two Gaussian mechanical quadratures formulae, important in theoretical as well as in practical applications:

Gauss's case:

$$\int_{-1}^{1} f(x)\,dx = \sum_{i=1}^{n} H_i\, f(x_i) + \left[\frac{n!}{1.3.5...(2n-1)}\right]^2 \frac{2}{(2n+1)!} f^{(2x)}(\eta), \tag{105}$$

$x_i$ being zeros of the Legendre polynomial of degree $n$.

Tchebycheff's case:

$$\int_{-1}^{1} f(x) \frac{dx}{\sqrt{(1-x^2)}} = \frac{\pi}{n} \sum_{i=1}^{n} f\left(\cos\frac{(2i-1)\pi}{2n}\right) + \frac{\pi}{2^{2n-1}(2n)!} f^{(2n)}(\eta). \tag{106}$$

## EXERCISES

1. Use a five-place table of trigonometric functions to find sin 37° as accurately as you can employing sin 30°, sin 40°, sin 50°,.... Discuss your error.

2. Find approximately the value of antilog 0·9763 452 given the table:

| $x$ | antilog $x$ |
|---|---|
| 0·95 | 8·912 509 |
| 0·96 | 9·120 108 |
| 0·97 | 9·332 543 |
| 0·98 | 9·549 926 |
| 0·99 | 9·772 372 |

Discuss the error.

3. From the table

| $x$ | $\sqrt{x}$ |
|---|---|
| 530·1 | 23·02 3901 |
| 540·1 | 23·24 0052 |
| 550·1 | 23·45 4211 |
| 560·1 | 23·66 6432 |

form a table of differences and calculate $\sqrt{530{\cdot}67459}$, $\sqrt{540{\cdot}67459}$, $\sqrt{550{\cdot}67459}$. In each case discuss the error.

4. Supply the values corresponding to $x = 0{\cdot}101$, $0{\cdot}103$, $0{\cdot}105$ in the following table:

| $x$ | $\sin x$ |
|---|---|
| 0·100 | 0·09983 3417 |
| 0·102 | 0·10182 3224 |
| 0·104 | 0·10381 2624 |
| 0·106 | 0·10580 1609 |

5. The following table gives values of the complete elliptic integral $E$ corresponding to values of $m$ ($= k^2$):

| $m$ | $E$ |
|---|---|
| 0·00 | 1·5707 96327 |
| 0·02 | 1·5629 12645 |
| 0·04 | 1·5549 68546 |
| 0·06 | 1·5469 62456 |
| 0·08 | 1·5388 92730 |

Insert the values corresponding to $m = 0{\cdot}01$, $0{\cdot}03$, $0{\cdot}05$, $0{\cdot}07$.

6. Find from the following data an approximate value of $\log 212$:

| | |
|---|---|
| $\log 210 = 2{\cdot}322\,2193$ | $\log 213 = 2{\cdot}328\,3796$ |
| $\log 211 = 2{\cdot}324\,2825$ | $\log 214 = 2{\cdot}330\,4138$ |

Discuss the error term.

7. From the following table of $\log\Gamma(n)$, determine approximately $\log\Gamma(1/2)$

| $n$ | $\log\Gamma(n)$ | $n$ | $\log\Gamma(n)$ |
|---|---|---|---|
| 2/12 | 0·74556 | 7/12 | 0·18432 |
| 3/12 | 0·55938 | 8/12 | 0·13165 |
| 4/12 | 0·42796 | 9/12 | 0·08828 |
| 5/12 | 0·32788 | 10/12 | 0·05261. |

8. Construct a polynomial of as low a degree as possible having the following values:

$$F(0) = 1, \quad F(1) = 2, \quad F(3) = -4, \quad F(4) = -3, \quad F(5) = 7.$$

9. Construct a polynomial of as low a degree as possible satisfying the following conditions:

$$F(0) = 1, \quad F'(0) = 2, \quad F(1) = 2, \quad F'(1) = 2, \quad F''(1) = 0, \quad F(2) = 3, \quad F'(2) = -1.$$

10. Construct a polynomial $F(x)$ of as low a degree as possible such that $F(x) = \log x$ and $F'(x) = 1/x$ at the points 1, 2, 3, 4, 5. Discuss the error.

11. Study the error for each of the trigonometric functions in turn if linear interpolation is used in an ordinary four-place table. Treat various sections of the table.

12. Write Newton's formula in terms of backward differences.

13. Deduce Newton's formula from Lagrange's formula.

14. Develop one formula which will include Stirling's, Bessel's, and Gauss's interpolation formulae as special cases.

15. Evaluate $\int_{-2}^{8} e^{-x^2}\,dx$, using a five-place table for $e^{-x^2}$, (1) by the trapezoid rule, $s = 10$; (2) by Simpson's rule, $s = 10$. Discuss the accuracy of the result in both cases.

16. Calculate Cotes' numbers, $n = 2, 3, 4, 5, 6$.

17. Calculate the abscissae in Tchebycheff's formula for mechanical quadratures, $n = 2, 3, 4, 5, 6$.

18. Calculate the abscissae and coefficients for Gauss's formula for mechanical quadratures ($p = 1$), $n = 2, 3, 4, 5, 6$.

19. Calculate the abscissae in the Tchebycheff-Gauss formula for mechanical quadratures, $n = 2, 3, 4, 5, 6$.

20. Use the Gauss abscissae and coefficients, $n = 5$, to calculate $\int_{-1}^{1} \sqrt{(1+x^3)}\, dx$. Discuss the error.

21. Use the Tchebycheff-Gauss abscissae and coefficients to calculate $\int^{1} \sqrt{(1+x^3)}\, dx$. Discuss the error.

# VII

# THE ELEMENTARY THEORY OF THE LINEAR RECURRENT RELATION

## 1. General discussion

An equation of the form

$$\phi\{i, y(i), y(i+1),..., y(i+n)\} = 0, \tag{1}$$

where $i$ is an integer or zero, is known as a recurrent relation, or difference equation with integral independent variable. Other terms sometimes used are recurrence formula or recursion formula. The letter $i$ rather than $x$ is used for the independent variable to emphasize its integral character.

Due to the relation

$$y(i+m) = \sum_{j=0}^{m} {}_mC_j \Delta^j y(i),$$

an equation such as (1) can also be written

$$F\{i, \Delta y(i),..., \Delta^n y(i)\} = 0. \tag{1'}$$

Examples of recurrent relations (difference equations) are

$$y(i+2)+2iy(i+1)+y(i) = 0,$$
$$\Delta^2 y(i)+4\Delta y(i)+4y(i) = 0.$$

These two equations are said to be linear.

*In general, equations of the form* (2) *or* (3) *below are called linear.*

$$p_0(i)y(i+n)+p_1(i)y(i+n-1)+...+p_n(i)y(i) = r(i), \tag{2}$$
$$P_0(i)\Delta^n y(i)+P_1(i)\Delta^{n-1}y(i)+...+P_n(i)y(i) = R(i), \tag{3}$$

where $p_0(i)$, $p_1(i)$,..., $p_n(i)$, $r(i)$ and likewise $P_0(i)$, $P_1(i)$,..., $P_n(i)$, $R(i)$ are defined over a set of integral values of $i$ such that

$$a \leqslant i \leqslant b. \tag{4}$$

*Equation* (2) *is of order n over* (4) *if and only if* $p_0(i).p_n(i) \neq 0$ *at any point of* (4).

Next consider

$$\Delta^3 y(i)+\Delta^2 y(i)-\Delta y(i)-y(i) = 0.$$

This reduces to $\quad y(i+3)-2y(i+2) = 0.$

We agree to call this a first-order equation. In fact, letting

$y(i+2) = z(i)$, we write it as $z(i+1)-2z(i) = 0$. Contrariwise, if (2) is reduced to the form (3) the term $p_0(i)\Delta^n y(i)$ will occur and no higher difference will occur. Hence, it is more convenient to define order of a difference equation by means of (2) as we have done. More specifically: *An equation of the form* (3) *is of the n-th order if and only if it is of the n-th order when written in the form* (2).

Inasmuch as $p_0(i) \neq 0$, we shall assume as our fundamental linear type

$$y(i+n)+p_1(i)y(i+n-1)+...+p_n(i)y(i) = r(i), \tag{5}$$

$p_n(i) \neq 0$ at any point.

By a solution of (5) over (4) we mean a function of $i$ satisfying (5) at all points of (4). The variable $i$, we recall, takes on integral values only. Such a function must be defined when $a \leqslant i \leqslant b+n$ on account of the occurrence of $y(i+n)$ in the equation. Inasmuch as there may be more than one solution, we shall speak of a *particular solution* in contrast to the *general solution*, by which we mean a formula which includes all solutions as special cases.

## 2. Homogeneous and non-homogeneous equations

Equation (5) is called *homogeneous* if $r(i) \equiv 0$. It is called *non-homogeneous* in the contrary case. We rewrite these equations here for convenience of reference:

$$y(i+n)+p_1(i)y(i+n-1)+...+p_n(i)y(i) = 0, \tag{6}$$

$$y(i+n)+p_1(i)y(i+n-1)+...+p_n(i)y(i) = r(i), \tag{7}$$

$$p_n(i) \neq 0, \qquad r(i) \not\equiv 0.$$

We assume that $p_1(i)$, $p_2(i)$,..., $p_n(i)$ are the same in both equations. A few theorems are immediate.

I. *If $y(i)$ is a solution of* (6) *so is $cy(i)$.*

II. *If $y_1(i)$ and $y_2(i)$ are solutions of* (6) *so is $y_1(i)+y_2(i)$.*

III. *If $y_1(i)$ is a solution of* (6) *and $Y(i)$ a solution of* (7), *then $y_1(i)+Y(i)$ is a solution of* (7).

Of a slightly different type is the following existence theorem.

THEOREM IV. *There exists one, and only one, solution of* (7) [(6)] *for which* $y(k) = a_0$, $y(k+1) = a_1, \ldots, y(k+n-1) = a_{n-1}$, *where* $a_0, a_1, \ldots, a_{n-1}$ *are preassigned arbitrarily and* $k$ *is an integral constant,* $a \leqslant k \leqslant b$.

This follows immediately if we remark that $y(a), \ldots, y(a+n-1)$ determine uniquely $y(a+n)$, that $y(a+1), \ldots, y(a+n)$ determine uniquely $y(a+n+1)$, etc. If instead of $a$ we use $k$, other than the initial point of (4), we can proceed similarly determining $y(k-1)$, etc., as necessary.

## 3. Linear equations of the first order

The general homogeneous linear recurrent relation of the first order is

$$y(i+1)-A(i)y(i) = 0, \tag{8}$$

where $A(i)$ is defined over (4) and is different from zero at all points.

The general solution of (8) is

$$y(i) = y(a)A(a)A(a+1)\ldots A(i-1). \tag{9}$$

It may be possible to express this product in closed form. Thus, in general,

$$\log y(i) = \log y(a)+\log A(a)+\log A(a+1)+\ldots+\log A(i-1).$$

This is a form to which summation processes may be applicable.

The general non-homogeneous equation of the first order is

$$y(i+1)-A(i)y(i) = B(i), \tag{10}$$

where $A(i) \neq 0$ and $B(i) \not\equiv 0$ are defined over (4). Assume $A(i)$ the same as in equation (8). To solve (10) we assume $u(i) \not\equiv 0$, a solution of (8), and strive to determine $v(i)$ so that $u(i).v(i)$ shall be a solution of (10). Substitute this in (10) and we find the following equation:

$$v(i)[u(i+1)-A(i)u(i)]+u(i+1)\Delta v(i) = B(i).$$

Inasmuch as $u(i)$ is a solution of (8),

$$u(i+1)\Delta v(i) = B(i).$$

We note that $u(i+1) \neq 0$ at any point. Otherwise $u(i) \equiv 0$, which contradicts our hypothesis. Hence

$$v(i) = \sum \frac{B(i)}{u(i+1)}.$$

We can use a definite sum if we like and write

$$y(i) = u(i)\left[\sum_{i=a}^{i-1} \frac{B(i)}{u(i+1)} + C\right].$$

This is a solution, as is shown by retracing steps. With the convention $\sum_{i=a}^{a-1} \frac{B(i)}{u(i+1)} = 0$, we have $C = \frac{y(a)}{u(a)}$. We note that $u(a) \neq 0$.

Example 1. $\qquad y(i+1)-3y(i) = e^i.$

Take $u(i) = 3^i$. Then

$$\begin{aligned} y(i) &= 3^i\left[\tfrac{1}{3}\sum_{i=0}^{i-1} (\tfrac{1}{3}e)^i + C\right] \\ &= 3^i\left[\tfrac{1}{3}\sum_{i=0}^{i-1} (\tfrac{1}{3}e)^i\right] + y(0)3^i \\ &= 3^{i-1}\left[\frac{(\frac{1}{3}e)^i-1}{\frac{1}{3}e-1}\right] + y(0)3^i \\ &= \frac{e^i-3^i}{e-3} + y(0)3^i. \end{aligned}$$

This is the general solution according to Theorem IV since $y(0)$ is arbitrary.

Example 2.

$$y(i+1) + \frac{3i+1}{3i+7}y(i) = \frac{i}{(3i+4)(3i+7)}.$$

Let

$$u(i) = (-1)^i \frac{1}{7}\cdot\frac{4}{10}\cdot\frac{7}{13}\cdots\frac{3i-2}{3i+4} = (-1)^i \frac{1.4}{(3i+1)(3i+4)}.$$

Then $$\frac{B(i)}{u(i+1)} = -\tfrac{1}{4}(-1)^i i.$$

Moreover,

$$\sum (-1)^i i = -\tfrac{1}{2}i(-1)^i - \tfrac{1}{2}\sum (-1)^i = (-1)^i[-\tfrac{1}{2}i+\tfrac{1}{4}]+C.$$

Hence, the general solution is given by

$$y(i) = \frac{-1}{(3i+1)(3i+4)}\left[-\tfrac{1}{2}i+\tfrac{1}{4}+(-1)^i C\right].$$

## 4. Fundamental systems of solutions

Let us consider equation (6). We know that if $y_1(i)$, $y_2(i)$,..., $y_n(i)$ are solutions then $C_1 y_1(i)+C_2 y_2(i)+\ldots+C_n y_n(i)$ is also a solution.

Definition. *The solutions* $y_1(i)$, $y_2(i)$,..., $y_m(i)$ of (6) *are said to be linearly dependent at all points of the interval* (4″)

$$a \leqslant i \leqslant b+n, \tag{4″}$$

*if there exist constants* $C_1$, $C_2$,..., $C_n$ *not all zero such that*

$$C_1 y_1(i)+C_2 y_2(i)+\ldots+C_m y_m(i) = 0$$

*at all points of* (4″).

Let

$$W(i) \equiv \begin{vmatrix} y_1(i) & y_2(i) & \cdots & y_n(i) \\ y_1(i+1) & y_2(i+1) & \cdots & y_n(i+1) \\ \cdots & \cdots & \cdots & \cdots \\ y_1(i+n-1) & y_2(i+n-1) & \cdots & y_n(i+n-1) \end{vmatrix}.$$

Theorem V. *A necessary and sufficient condition that n solutions of* (6) *be linearly dependent over* (4″) *is* $W(a) = 0$.

First: The fact that the condition is necessary is immediate from Cramer's rule in the theory of linear algebraic equations.

Second: We shall prove the condition sufficient.

Write the following equations:

$$\begin{gathered} C_1 y_1(a)+C_2 y_2(a)+\ldots+C_n y_n(a) = 0, \\ C_1 y_1(a+1)+C_2 y_2(a+1)+\ldots+C_n y_n(a+1) = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ C_1 y_1(a+n-1)+C_2 y_2(a+n-1)+\ldots+C_n y_n(a+n-1) = 0. \end{gathered} \tag{11}$$

We can determine the $C$'s not all zero so as to satisfy these equations. Moreover, if equations (11) are satisfied, then so is

$$C_1 y_1(a+n)+C_2 y_2(a+n)+\ldots+C_n y_n(a+n) = 0,$$

since, by (6), the left-hand member is a linear combination of the left-hand members of (11). In a similar way

$$C_1 y_1(a+n+1)+C_2 y_2(a+n+1)+\ldots+C_n y_n(a+n+1) = 0;$$

and then by induction

$$C_1 y_1(i)+C_2 y_2(i)+\ldots+C_n y_n(i) = 0$$

so long as $a \leqslant i \leqslant b+n$. In other words we have proved linear dependence over (4″) and have completed our theorem.

We now prove the relation

$$W(i+1) = (-1)^n p_n(i) W(i). \tag{12}$$

To do this write out $W(i+1)$ and substitute for $y_1(i+n),\ldots, y_n(i+n)$ from (6).

We immediately see that if $W(a) = 0$, $W(i) \equiv 0$ over (4). As a matter of fact if $W(i) = 0$ at any part of (4) $W(i) \equiv 0$ over (4). We conclude the following theorem.

THEOREM VI. *A necessary and sufficient condition that* $y_1(i)$, $y_2(i),\ldots, y_n(i)$ *be linearly dependent over* (4″) *is* $W(i) \equiv 0$ *over*

$$a \leqslant i \leqslant b+1. \tag{4'}$$

DEFINITION. *The determinant* $W(i)$ *is called the Wronskian of the set of solutions of* (6), $y_1(i)$, $y_2(i),\ldots, y_n(i)$.

THEOREM VII. *Every* $n+1$ *solutions of* (6) *are linearly dependent over* (4″).

This is readily proved as follows: Any solution is completely determined by its values at $a$, $a+1,\ldots, a+n-1$. Now the equations

$$C_1 y_1(a)+C_2 y_2(a)+\ldots+C_{n+1} y_{n+1}(a) = 0,$$
$$C_1 y_1(a+1)+C_2 y_2(a+1)+\ldots+C_{n+1} y_{n+1}(a+1) = 0,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$C_1 y_1(a+n-1)+C_2 y_2(a+n-1)+\ldots+C_{n+1} y_{n+1}(a+n-1) = 0$$

can always be satisfied by $C$'s not all zero. As in the case where we were considering $n$ solutions only, we show that with this determination of the $C$'s

$$C_1 y_1(i)+C_2 y_2(i)+\ldots+C_{n+1} y_{n+1}(i) \equiv 0, \quad a \leqslant i \leqslant b+n,$$

thus establishing the theorem.

DEFINITION. *A set of solutions of* (6), $y_1(i),\ldots, y_m(i)$, *is called a fundamental system of solutions if every solution is expressible in the form* $C_1 y_1(i)+\ldots+C_m y_m(i)$, *and if this is true of no sub-set of the system.*

THEOREM VIII. *A set of solutions* $y_1(i),\ldots, y_m(i)$, $m < n$, *never form a fundamental system.*

To prove this consider the matrix

$$\begin{pmatrix} y_1(a) & y_2(a) & \cdots & y_m(a) \\ y_1(a+1) & y_2(a+1) & \cdots & y_m(a+1) \\ \cdot & \cdot & \cdot & \cdot \\ y_1(a+n-1) & y_2(a+n-1) & \cdots & y_m(a+n-1) \end{pmatrix}.$$

Suppose this of rank $r \leqslant m$ and suppose that a non-vanishing determinant of order $r$ lies in the upper left-hand corner. Consider the particular solution $\bar{y}(i)$, where

$$\bar{y}(a) = \ldots = \bar{y}(a+n-2) = 0, \qquad \bar{y}(a+n-1) = 1.$$

Set up the equations

$$C_1 y_1(a)+C_2 y_2(a)+\ldots+C_m y_m(a) = \bar{y}(a),$$
$$C_1 y_1(a+1)+C_2 y_2(a+1)+\ldots+C_m y_m(a+1) = \bar{y}(a+1),$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$C_1 y_1(a+n-1)+C_2 y_2(a+n-1)+\ldots+C_m y_m(a+n-1) = \bar{y}(a+n-1).$$

Consider these as equations in $C_1,\ldots, C_m$. The matrix of the coefficients of this system of equations is of rank $r$. The augmented matrix, namely the matrix formed by adding the column of right-hand members to the matrix of the coefficients, contains the determinant

$$\begin{vmatrix} y_1(a) & y_2(a) & \cdots & y_r(a) & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ y_1(a+r) & y_2(a+r) & \cdots & y_r(a+r) & 0 \\ y_1(a+n-1) & y_2(a+n-1) & \cdots & y_r(a+n-1) & 1 \end{vmatrix} \neq 0$$

and hence is of rank at least $r+1$. Consequently the system of equations is inconsistent. If the determinant of order $r$ in the upper left-hand corner is zero, simply choose $y$ with initial values all zero but one and locate this non-zero value so that

the augmented matrix contains a determinant of order $r+1$ like the above.

The inconsistency of the equations proves our theorem.

THEOREM IX. *No $n+1$ solutions form a fundamental system.*

This is an immediate consequence of Theorem VII.

THEOREM X. *A necessary and sufficient condition that $y_1(i),\ldots,$ $y_n(i)$ form a fundamental system of solutions of* (6) *is that they be linearly independent.*

First, to prove the condition sufficient we let $Y(i)$ be any solution and write

$$C_1 y_1(a)+C_2 y_2(a)+\ldots+C_n y_n(a) = Y(a),$$
$$C_1 y_1(a+1)+C_2 y_2(a+1)+\ldots+C_n y_n(a+1) = Y(a+1),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$C_1 y_1(a+n-1)+C_2 y_2(a+n-1)+\ldots+C_n y_n(a+n-1)$$
$$= Y(a+n-1).$$

These equations can be solved for the $C$'s which then serve to express $Y(i)$ linearly in terms of $y_1(i),\ldots, y_n(i)$.

To prove the condition necessary it suffices to remark that if the solutions $y_1(i),\ldots, y_n(i)$ constitute a fundamental system and if they are not linearly independent, at least one of them can be expressed linearly in terms of the others and hence any solution can be expressed as a linear combination of a sub-set $y_1(i),\ldots, y_n(i)$. This contradicts the hypothesis that they constitute a fundamental system.

THEOREM XI. *A fundamental system of solutions of* (6) *consists of exactly n linearly independent solutions.*

This is an immediate consequence of the last three theorems. It is well to remark that *fundamental systems of solutions exist.* We can assume initial values so that $W(a) \neq 0$.

THEOREM XII. *Any n functions $y_1(i),\ldots, y_n(i)$, which are defined for $a \leqslant i \leqslant b+n$ and are such that $W(i) \neq 0$ at any point of* (4), *determine uniquely a homogeneous, linear recurrent relation of the n-th order with coefficients defined over* (4), *for which $y_1(i), y_2(i),\ldots, y_n(i)$ constitute a fundamental system of solutions.*

We not only prove the theorem but develop formulae for the coefficients of the required equation, of the form (6), if we substitute in (6) successively and solve the resulting equations for $p_1(i),..., p_n(i)$. We note that $p_n(i) = -\frac{W(i+1)}{W(i)} \neq 0$ so that the equation is actually of order $n$.

THEOREM XIII. *The general solution of* (6) *is given by*

$$y(i) = C_1 y_1(i)+C_2 y_2(i)+...+C_n y_n(i), \tag{13}$$

*where* $y_1(i),..., y_n(i)$ *constitute a fundamental system of solutions and* $C_1,..., C_n$ *are constants.*

This follows directly from the definition of a fundamental system of solutions and Theorem XI and the fact that a fundamental system of solutions exists.

## 5. Non-homogeneous equations

THEOREM XIV. *If* $Y(i)$ *is a particular solution of* (7) *and* $y_1(i), y_2(i),..., y_n(i)$ *are a fundamental system of solutions of* (6) *then the general solution of* (7) *is given by the formula*

$$y(i) = C_1 y_1(i)+C_2 y_2(i)+...+C_n y_n(i)+Y(i). \tag{14}$$

This is true since the constants $C_1,..., C_n$ can be determined so that $y(a), y(a+1),..., y(a+n-1)$ have arbitrary preassigned values. The term 'complementary function' is used to describe the expression $C_1 y_1(i)+C_2 y_2(i)+...+C_n y_n(i)$ when speaking of the general solution of (7) as given by (14).

Our next problem is to develop methods of finding a particular solution $Y(i)$. We first develop below the so-called method of '*variations of constants*', which is similar to the method bearing the same name in the analogous theory of linear differential equations. Let us again consider (7), namely the non-homogeneous equation:

$$y(i+n)+p_1(i)y(i+n-1)+...+p_n(i)y(i) = r(i). \tag{7}$$

Suppose $y_1(i), y_2(i),..., y_n(i)$ form a fundamental system of solutions of the corresponding homogeneous equation (6). We wish to determine functions $C_1(i), C_2(i),..., C_n(i)$ so that

$$y(i) = C_1(i)y_1(i)+C_2(i)y_2(i)+...+C_n(i)y_n(i) \tag{15}$$

satisfies (7). We write down

$$\begin{aligned}
y(i+1) &= C_1(i)y_1(i+1)+C_2(i)y_2(i+1)+\dots+C_n(i)y_n(i+1),\\
y(i+2) &= C_1(i)y_1(i+2)+C_2(i)y_2(i+2)+\dots+C_n(i)y_n(i+2),\\
&\cdots\\
y(i+n-1) &= C_1(i)y_1(i+n-1)+C_2(i)y_2(i+n-1)+\dots+\\
&\quad +C_n(i)y_n(i+n-1).
\end{aligned} \tag{16}$$

Since $C(i+1) = C(i)+\Delta C(i)$, equations (16) are true if the $(n-1)$ equations (17) below are satisfied:

$$\begin{aligned}
&y_1(i+1)\Delta C_1(i)+y_2(i+1)\Delta C_2(i)+\dots+y_n(i+1)\Delta C_n(i) = 0,\\
&y_1(i+2)\Delta C_1(i)+y_2(i+2)\Delta C_2(i)+\dots+y_n(i+2)\Delta C_n(i) = 0,\\
&\cdots\\
&y_1(i+n-1)\Delta C_1(i)+y_2(i+n-1)\Delta C_2(i)+\dots+y_n(i+n-1)\Delta C_n(i)\\
&\qquad = 0.
\end{aligned} \tag{17}$$

Moreover, if equations (16) are satisfied, then

$$\begin{aligned}
y(i+n) &= C_1(i)y_1(i+n)+C_2(i)y_2(i+n)+\dots+C_n(i)y_n(i+n)+\\
&\quad +y_1(i+n)\Delta C_1(i)+y_2(i+n)\Delta C_2(i)+\dots+y_n(i+n)\Delta C_n(i).
\end{aligned} \tag{18}$$

Substitute (16) and (18) in the given equation. We get

$$y_1(i+n)\Delta C_1(i)+y_2(i+n)\Delta C_2(i)+\dots+y_n(i+n)\Delta C_n(i) = r(i). \tag{19}$$

Add this to the set of equations (17) and solve the resulting $n$ equations for $\Delta C_1(i), \Delta C_2(i),\dots, \Delta C_n(i)$. We find

$$\Delta C_1(i) = \frac{\begin{vmatrix} 0 & y_1(i+1) & . & . & . & y_1(i+n) \\ 0 & y_2(i+1) & . & . & . & y_2(i+n) \\ . & . & . & . & . & . \\ r(i) & y_n(i+1) & . & . & . & y_n(i+n) \end{vmatrix}}{W(i+1)} \equiv \frac{\delta_{1n}}{W(i+1)}$$

with precisely similar formulae for $\Delta C_2(i),\dots, \Delta C_n(i)$. We know that $W(i+1) \neq 0$, and hence that the solution is possible and unique. We now get

$$C_j(i) = \sum \frac{\delta_{jn}}{W(i+1)}, \quad j = 1, 2,\dots, n.$$

We substitute these functions in (15). That we actually have obtained a solution is shown by retracing our steps or by substitution in (7). As a matter of practical procedure the general solution of (7) is obtained by particularizing the sums appearing in the formulae for $C_1(i),..., C_n(i)$, substituting in (15), and adding the complementary function.

## 6. Linear equations with constant coefficients

Let us consider

$$y(i+n)+A_1 y(i+n-1)+...+A_n y(i) = 0, \tag{20}$$

where $A_1,..., A_n$ are constants and $A_n \neq 0$. This can conveniently be written, employing the symbol $E$ introduced in Chapter I,

$$\phi(E)y(i) = 0,$$

where $\phi(x) = x^n+A_1 x^{n-1}+...+A_n$.

Similarly, if the equation written in $\Delta$-form is

$$\Delta^n y(i)+B_1 \Delta^{n-1}y(i)+...+B_n y(i) = 0$$

we can write

$$\chi(\Delta)y(i) = 0,$$

where $\chi(x) = x^n+B_1 x^{n-1}+...+B_n$.

Let us substitute $y(i) = \alpha^i$ in (20). We find the following necessary condition that $\alpha^i$ be a solution of (20):

$$\phi(\alpha) = \alpha^n+A_1 \alpha^{n-1}+...+A_n = 0. \tag{21}$$

We call this the *auxiliary equation* of (20). Conversely, if $\alpha$ is a root of this equation, $\alpha^i$ is a solution, as is immediately proved by trial. If the roots of (21) are all distinct, $\alpha_1$, $\alpha_2$,..., $\alpha_n$, then $\alpha_1^i$, $\alpha_2^i$,..., $\alpha_n^i$ constitute a fundamental system of solutions:

$$W(i) = \alpha_1^i \alpha_2^i \dots \alpha_n^i(\alpha_1-\alpha_2)...(\alpha_{n-1}-\alpha_n) \neq 0.$$

It remains to consider multiple roots of the auxiliary equation (21).

If $\alpha_1 = \alpha_2$ consider $i\alpha_1^i$, and substitute in (20). We find that a necessary and sufficient condition that $i\alpha_1^i$ be a solution of (20) is

$$\phi'(\alpha_1) = 0,$$

which is precisely the condition that $\alpha_1$ be a double root of (21). Similarly, if $\alpha_1 = \alpha_2 = ... = \alpha_m$ then $\alpha_1^i$, $i\alpha_1^i$,..., $i^{m-1}\alpha_1^i$ are all

solutions of (20). Treating each multiple root in this way we obtain a *fundamental* system of solutions as is proved by the Wronskian test.

In case $A_1, A_2,\ldots, A_n$ are real, imaginary roots occur in pairs, say $\gamma\pm\delta\sqrt{(-1)}$, and the two solutions $\{\gamma+\delta\sqrt{(-1)}\}^i$ and $\{\gamma-\delta\sqrt{(-1)}\}^i$ can be replaced by linear combinations, namely $\rho^i\cos i\theta$ and $\rho^i\sin i\theta$, where

$$\gamma\pm\delta\sqrt{(-1)} = \rho\{\cos\theta\pm\sqrt{(-1)}\sin\theta\}.$$

As an example, consider the equation

$$y(i+2)-7y(i+1)+6y(i) = i. \tag{22}$$

Here $\qquad \phi(\alpha) = \alpha^2-7\alpha+6, \qquad \alpha_1 = 6, \qquad \alpha_2 = 1.$

Thus, the complementary function is $C_1\,1^i+C_2\,6^i = C_1+C_2\,6^i$.

We next find a particular solution of (22) by the method of variation of constants. This leads to

$$\Delta C_1(i) = -\frac{i}{5}, \qquad \Delta C_2(i) = \frac{1}{30}i\left(\frac{1}{6}\right)^i,$$

$$C_1(i) = -\frac{1}{10}i(i-1), \qquad C_2(i) = -\frac{1}{25}i\left(\frac{1}{6}\right)^i-\frac{1}{125}\left(\frac{1}{6}\right)^i.$$

Whereupon the general solution of (22) is given by

$$y(i) = C_1+C_2\,6^i-\frac{1}{10}i^2+\frac{3}{50}i-\frac{1}{125} = C_1+C_2\,6^i-\frac{1}{10}i^2+\frac{3}{50}i.$$

## 7. The method of undetermined coefficients

The determination of a particular solution by 'undetermined coefficients' is similar to the analogous method for differential equations. As an illustration we exhibit this method on the example just treated, namely equation (22). We write this in difference form:

$$\Delta^2y(i)-5\Delta y(i) = i. \tag{22'}$$

Assume a particular solution in the form of a polynomial of the second degree, namely

$$Y(i) = A\frac{i(i-1)}{2}+Bi, \quad A,\ B \text{ constants.}$$

We substitute this in (22′) and readily find $A = -\frac{1}{5}$, $B = -\frac{1}{25}$.
We easily verify that $-\frac{1}{5}\frac{i(i-1)}{2}-\frac{1}{25}i$ is a solution.

In general, suppose that we are given the linear non-homogeneous equation with constant coefficients

$$\phi(E)y(i) = (E^n+A_1E^{n-1}+\ldots+A_n)y(i) = Cm^i i^{(k)}, \qquad (23)$$

where $\phi(E)$ is of degree $n$ and, as we recall,

$$Eu(i) = (1+\Delta)u(i) = u(i+1).$$

Suppose that $m$ is an $r$-fold root of $\phi(\alpha) = 0$ ($r$ may be zero).

We first remark the formula

$$\phi(E)m^i u(i) = m^i\phi(m+m\Delta)u(i)$$

or, written in another way by means of Taylor's series,

$$\phi(E)m^i u(i) = m^i\left[\phi(m)+\phi'(m)m\Delta+\ldots+\frac{\phi^{(r)}(m)}{r!}m^r\Delta^r+\ldots+\right.$$
$$\left.+\frac{\phi^{(n)}(m)}{n!}m^n\Delta^n\right]u(i). \qquad (24)$$

This formula follows from the fact that

$$E^j\{m^i u(i)\} = (E^j m^i)(1+\Delta)^j u(i) = m^i\{m^j(1+\Delta)^j u(i)\}$$
$$= m^i(m+m\Delta)^j u(i)$$

and the fact that $\phi(E)$ is a polynomial.

We next assume a particular solution as follows:

$$Y(i) = (a_1 i^{(k+r)}+a_2 i^{(k+r-1)}+\ldots+a_k i^{(r)})m^i, \qquad (25)$$

where $a_1, a_2,\ldots, a_k$ are unknown constants.

Substitute $Y(i)$ as given by (25) in (23) using (24). Note that

$$\phi(m) = \phi'(m) = \ldots = \phi^{(r-1)}(m) = 0.$$

We get

$$\phi(E)Y(i) = Cm^i\left[i^{(k)}a_1\frac{\phi^{(r)}(m)}{r!}m^r(k+r)\ldots(k+1)+\right.$$
$$+i^{(k-1)}\left\{a_1B_{11}+a_2\frac{\phi^{(r)}(m)}{r!}m^r(k+r-1)\ldots k\right\}+$$
$$+i^{(k-2)}\left\{a_1B_{21}+a_2B_{22}+a_3\frac{\phi^{(r)}(m)}{r!}m^r(k+r-2)\ldots(k-1)\right\}+$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\left.+i^{(r)}\left\{a_1B_{k1}+a_2B_{k2}+\ldots+a_{k+1}\frac{\phi^{(r)}(m)}{r!}m^r(r!)\right\}\right] = Cm^i i^{(k)}. \qquad (26)$$

Here the $B$'s result from operating with $\Delta^{r+1},\ldots, \Delta^n$.

Now $\phi^{(r)}(m) \neq 0$ as $m$ is only an $r$-fold root of $\phi(\alpha) = 0$. Moreover $m \neq 0$ since we are assuming that (23) is non-homogeneous. Also each of the products

$$(k+r)...(k+1), \quad (k+r-1)...(k-1), \quad ..., \quad r!$$

is greater than zero. It is immediate that we can equate coefficients of $i^{(k)}$, $i^{(k-1)}$,..., 1 on both sides of (26) and solve successively for $a_1, a_2,..., a_{k+1}$. The determination is such as to make $Y(i)$ a solution of (23).

In case there are several terms of the form $Cm^i i^{(k)}$, each can be treated separately, the work being carried forward in any convenient way.

If the right-hand member of (23) is of the form $Cm^i i^{(k)}\cos pi$, it can be replaced by $\frac{1}{2}CM^i i^{(k)}+\frac{1}{2}C\bar{M}^i i^{(k)}$, where

$$M = m\{\cos p+\sqrt{(-1)}\sin p\}$$

and $\bar{M}$ is the conjugate of $M$. If all coefficients in $\phi(\alpha)$ are real, then if $M$ is an $r$-fold root so is $\bar{M}$. If $C$ is also real, a moment's reflection shows that $\bar{a}_1,..., \bar{a}_{k+1}$ corresponding to $\bar{M}^i$ are conjugate to $a_1,..., a_{k+1}$ corresponding to $M^i$. As a consequence if $Y(i)$ denotes the sum of the particular solutions obtained corresponding to $\frac{1}{2}CM^i i^{(k)}$ and $\frac{1}{2}C\bar{M}^i i^{(k)}$ respectively, then $Y(i)$ is real.

If the right-hand member contains $Cm^i i^{(k)}\sin pi$ simply replace it by $\frac{1}{2\sqrt{(-1)}}CM^i i^{(k)}-\frac{1}{2\sqrt{(-1)}}C\bar{M}^i i^{(k)}$ and proceed as previously.

## 8. The method of operators

Consider the equation

$$\phi_1(E)y \equiv (E-\alpha)y = p(i). \tag{27}$$

We denote the solution of this equation by

$$y(i) = (E-\alpha)^{-1}p(i).$$

Consider now the equation

$$\phi_2(E)y = (E-\alpha_1)(E-\alpha_2)y = p(i). \tag{28}$$

Then $$y(i) = (E-\alpha_2)^{-1}[(E-\alpha_1)^{-1}p(i)],$$

where the negative exponent means the solution of the corresponding equation of the first order. That $y$ which is thus given

is truly a solution of (28) follows from the properties of the operator $E$. In general if the equation has the form

$$\phi_n(E)y \equiv (E-\alpha_1)(E-\alpha_2)...(E-\alpha_n)y = p(i)$$

a solution is given by

$$y(i) = (E-\alpha_n)^{-1}[(E-\alpha_{n-1})^{-1}\{...(E-\alpha_1)^{-1}\}]p(i),$$

that is, by the successive solution of $n$ first-order equations. The general solution of (28) is obtained if the general solution of each first-order equation is obtained and subsequently used. It is, however, more convenient to particularize the solution at each step on account of the ease of obtaining the complementary function.

It may happen that the equation is easier to factor in the difference form. Thus we may have

$$F(\Delta)y \equiv (\Delta-\beta_1)(\Delta-\beta_2)...(\Delta-\beta_n)y = p(i).$$

Procedure is exactly as formerly, namely,

$$y(i) = (\Delta-\beta_n)^{-1}(\Delta-\beta_{n-1})^{-1}...(\Delta-\beta_1)^{-1}p(i)$$

will give the general solution if the general solutions of all first-order equations involved are obtained. However, as before, the best procedure is to find a particular solution only, particularizing as successive first-order equations are solved.

## 9. The linear equation whose coefficients are power series in a parameter

For compactness in writing we shall confine our attention to the equation of the second order.

Consider

$$y(i+2)+p(i,\mu)y(i+1)+q(i,\mu)y(i) = R(i,\mu), \tag{29}$$

where

$$p(i,\mu) = p_0(i)+p_1(i)\mu+p_2(i)\mu^2+..., \tag{30}$$

$$q(i,\mu) = q_0(i)+q_1(i)\mu+q_2(i)\mu^2+..., \tag{31}$$

$$R(i,\mu) = R_0(i)+R_1(i)\mu+R_2(i)\mu^2+.... \tag{32}$$

We assume series (30), (31), and (32) to be convergent when

$$|\mu| < \mu_0 \quad \text{and} \quad a \leqslant i \leqslant b.$$

THEOREM XV. *A solution of* (29) *satisfying initial conditions* $y(a) = A$, $y(a+1) = B$ *can be written*

$$y(i) = y_0(i)+y_1(i)\mu+y_2(i)\mu^2+\dots, \tag{33}$$

*which series is convergent,* $|\mu| < \mu_0$ *and* $a \leqslant i \leqslant b+2$. *The functions* $y_0(i)$, $y_1(i)$,... *are uniquely determined and satisfy the following equations:*

$$y_n(i+2)+p_0(i)y_n(i+1)+q_0(i)y_n(i)+X_n(i) = 0, \tag{34}$$

*where*

$$X_n(i) = \sum_{\mu=0}^{n-1}\{y_\mu(i+1)p_{n-\mu}(i)+y_\mu(i)q_{n-\mu}(i)\}+R_n(i), \quad n > 0,$$

$$X_0(i) = R_0(i),$$

$$y_0(a) = A, \qquad y_0(a+1) = B,$$

$$y_j(a) = y_j(a+1) = 0, \quad j > 0.$$

This theorem is particularly useful in determining the functions $y_0(i)$, $y_1(i)$,..., if $p_0(i)$, $q_0(i)$ are constants. The successive solving of (34) for these functions then falls directly under the methods developed in this chapter. We note that the reduced equation in each instance is the same.

We proceed to the proof of the theorem. The proof is a sequel to well-known theorems on power series. Let $i = a$. Substitute for $p$, $q$, and $R$ from (30), (31), (32) and substitute $A$ for $y(a)$ and $B$ for $y(a+1)$. From this equation we determine $y(a+2)$ as a power series in $\mu$. We then determine $y(a+3)$ then $y(a+4)$,..., $y(b+2)$. If the series for $p$, $q$, and $R$ converge for all values of $i$ we can proceed indefinitely.

Two developments of the form (33) are not possible, for if

$$y_0(i)+y_1(i)\mu+y_2(i)\mu^2+\dots = \bar{y}_0(i)+\bar{y}_1(i)\mu+\bar{y}_2(i)\mu^2+\dots$$

when $|\mu| < \mu_0$ and $a \leqslant i \leqslant b+2$ then

$$y_0(i) = \bar{y}_0(i), \qquad y_1(i) = \bar{y}_1(i), \qquad \dots.$$

In order to calculate the functions $y_0(i)$, $y_1(i)$,... we substitute (30), (31), (32), (33) in (29) and equate coefficients of $\mu^j$. We find

$$y_n(i+2)+p_0(i)y_n(i+1)+q_0(i)y_n(i)+X_n(i) = 0,$$

$$X_0(i) = R_0(i),$$

$$X_n(i) = \sum_{\mu=0}^{n-1}\{y_\mu(i+1)p_{n-\mu}(i)+y_\mu(i)q_{n-\mu}(i)\}+R_n(i),$$

$$n = 1, 2,\dots$$

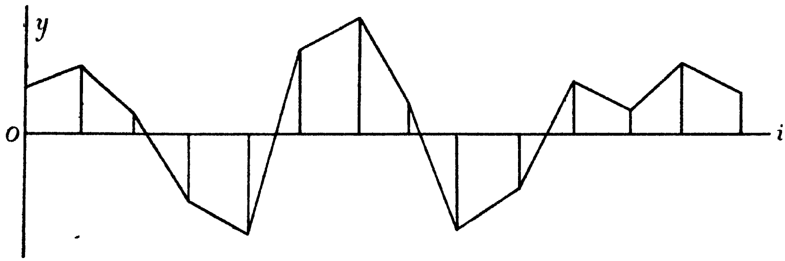
Solutions of these equations subject to the conditions

$$y_0(a) = A, \qquad y_0(a+1) = B, \qquad y_j(a) = y_j(a+1) = 0, \quad j > 0,$$

substituted in (33) form a series which satisfies (29) and the initial conditions.

## 10. Graphical representation of solutions of recurrent relations

A real solution of a recurrent relation can be plotted in the Cartesian plane as a set of isolated points. It is, however,




customary to connect these points with straight line segments forming a broken straight line which we will call the graph of the solution. Points where the function of the continuous variable defined by this graph is zero will be called 'nodes'. (See figure.)

## 11. The difference interval different from unity

In a recurrent relation as it actually occurs it is not necessary that the set of congruent points be the integers nor even that the difference interval be 1. This is only a matter of form. The relation

$$y\{a+(i+n)h\}+p_1(a+ih)y\{a+(i+n-1)h\}+\ldots+$$
$$+p_n(a+ih)y(a+ih) = r(a+ih)$$

can be written

$$Y(j+n)+P_1(j)Y(j+n-1)+\ldots+P_n(j)Y(j) = R(j),$$

which is of the type studied. Properties of $Y$ can be transferred to $y$ or vice versa simply by means of the relationship

$$Y(j+b) = y\{a+(i+b)h\}$$

where $b = 0, 1, \ldots$.

## EXERCISES

1. (*a*) The Fibonacci series is characterized as follows:

$$y(i) = y(i-1)+y(i-2).$$

Solve for $y(i)$, if $y(0) = 1$, $y(1) = 2$.

(*b*) Solve for $y(i)$ if

$$y(i) = 3y(i-1)+5y(i-2), \qquad y(0) = 2, \qquad y(1) = 3.$$

2. Find the general solution of the following equations. Use at least two methods for finding a particular solution in each case.

(*a*) $y(i+3)-6y(i+2)+11y(i+1)-6y(i) = 1+4^i$.

(*b*) $y(i+3)-8y(i) = 2^i+i2^i$.

(*c*) $(E-1)^2(E+1)^2y(i) = i^3+\cos i$.

(*d*) $(E-2)^3(E+1) = i^4 2^i-3i^2+1$.

3. Show how to make the finding of the general solution of a homogeneous linear equation depend upon the solution of a similar type of equation but of lesser order, given one solution not identically zero.

4. If $y_1(i)$, $y_2(i),\ldots,$ $y_n(i)$ are linearly independent solutions of the equation

$$y(i+n)+p_1(i)y(i+n-1)+\ldots+p_n(i)y(i) = 0,$$

where $p_n(i) \neq 0$ and $n$ is even, show that $W(i)$ retains a fixed sign. What is the corresponding theorem if $n$ is odd?

5. If $y_1(i)$ and $y_2(i)$ are linearly independent solutions of

$$y(i+2)+p_1(i)y(i+1)+p_2(i)y(i) = 0, \quad p_2(i) > 0,$$

show that nodes of $y_1$ and $y_2$ separate each other.

6. If

$$z(i)\{y(i+2)+p_1(i)y(i+1)+p_2(i)y(i)\} = \Delta\{\eta(i)y(i+1)+\zeta(i)y(i)\}$$

then $z(i)$ is called an integrating factor of

$$y(i+2)+p_1(i)y(i+1)+p_2(i)y(i) = 0.$$

Set up a linear recurrent relation that must be satisfied by $z(i)$.

7. Generalize Exercise 6.

8. Prove that it is possible to choose a function $\phi(i) \neq 0$ so that the substitution $y(i) = \phi(i)\bar{y}(i)$ transforms an equation

$$\Delta^2y(i)+p(i)\Delta y(i)+q(i)y(i) = 0,$$
$$-1-p(i)+q(i) \neq 0,$$

into an equation of the form

$$\Delta^2y(i)+Q(i)y(i+1) = 0.$$

9. Generalize Exercise 8.

VIII

# MAXIMA AND MINIMA OF FINITE SUMS

## 1. Maxima and minima of functions of more than one variable

WE call attention to some theorems relative to maxima and minima in order that the reader may have them freshly in mind. Given a function $f(x_1,\ldots, x_n)$ all of whose third derivatives exist in the neighbourhood in question:

(I) A necessary condition for a maximum or minimum is that

$$\frac{\partial f}{\partial x_i} = 0, \quad i = 1, 2,\ldots, n.$$

(II) If (I) is satisfied at a point $x_1^{(1)},\ldots, x_n^{(1)}$, then a sufficient condition for a minimum at that point is that the form

$$\left[\left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n}\right)^2 f(x_1,\ldots, x_n)\right]_{(x_1^{(1)},\ldots,x_n^{(1)})}$$

be positive definite. A sufficient condition for a maximum is that it be negative definite.

## 2. The finite sum

Consider the sum

$$\phi = \sum_{i=0}^{n-1} F(x_i, y, \Delta y), \tag{1}$$

where $y$ is to be determined as a function of $x_i$ so as to make $\phi$ a maximum or minimum, where $x_{i+1} > x_i$, and where

$$\Delta y = y(x_{i+1}) - y(x_i).$$

We shall assume at least the existence of all third derivatives of $F$ in all neighbourhoods considered.

Evidently $x_i$ is a function of $i$. An analytic form for this function can be written down in a variety of ways. A satisfactory way is by means of the Lagrange interpolation formula. Such a transformation on $x_i$ reduces (1) to the form

$$\phi = \sum_{i=0}^{n-1} f(i, y, p), \tag{2}$$

where now $y$ is to be determined as a function of $i$ and

$$p = \Delta y = y(i+1)-y(i).$$

The function $y(i)$ can in turn be written as a function of $x_i$ if desired; the Lagrange interpolation formula is again an adequate instrument. Partly on account of this transformation we shall consider only the sum (2), although the major reason is for simplicity in writing. The reader will have no difficulty in carrying through like reasoning for (1) if he so desires.

For convenience we shall usually replace $f(i, y, p)$ simply by $f(i)$.

Let us apply to (2) the first necessary condition for a minimum. We have

$$\frac{\partial\phi}{\partial y(i)} = f_y(i)-f_p(i)+f_p(i-1) = 0, \tag{3}$$

which can be written

$$f_y(i)-\Delta f_p(i-1) = 0.$$

This holds when $0 < i \leqslant n-1$. We can rewrite (3)

$$f_y(i+1)-\Delta f_p(i) = 0, \tag{4}$$

where now $0 \leqslant i \leqslant n-2$. This in turn can be written

$$\Delta[f_y(i)-f_p(i)]+f_y(i) = 0. \tag{4'}$$

THEOREM I. *The satisfaction of equation* (4) *is a necessary condition that* $\phi$ *be a maximum or minimum.*

In case $y(0)$ and/or $y(n)$ are variable as well as the $y$'s at intermediate points we have in addition the relations

$$f_y(0)-f_p(0) = 0$$

and/or

$$f_p(n-1) = 0. \tag{5}$$

These are boundary conditions which result from the fact that $f(-1)$ does not occur in the sum $\phi$ and consequently that the last term of (3) is lacking when $i = 0$ and/or that $f(n)$ does not occur in $\phi$.

The similarity of (4) to Euler's equation in the calculus of variations will be immediate to those who have studied that subject and it is interesting to derive it by the method of variations.

A function which satisfies (4) if $y(0)$ and $y(n)$ are fixed, or (4) and (5) in the contrary case, will be called a critical function.

The sufficiency condition, as quoted under (II), results as follows† when applied to $\phi$. Let $y(i)$ be a critical function and let us assume that $y(0)$ and $y(n)$ are constants. Let

$$a_{ij} = \frac{\partial^2\phi}{\partial y(i)\partial y(j)}.$$

Then the form in (II) can be written

$$\begin{aligned} P(h_1,\dots,h_n) = {} & a_{11}h_1^2+a_{12}h_1h_2+0+0+\dots+0+0+0+ \\ & +a_{21}h_1h_2+a_{22}h_2^2+a_{23}h_2h_3+0+\dots+0+0+0+ \\ & +0+a_{32}h_3h_2+a_{33}h_3^2+a_{34}h_3h_4+\dots+0+0+0+ \\ & +\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot + \\ & +0+0+0+0+\dots+a_{n,n-1}h_{n-1}h_n+a_{nn}h_n^2 \end{aligned} \tag{6}$$

with $a_{ij} = a_{ji}$.

Let

$$D_0 = 1, \quad D_1 = a_{11}, \quad D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \dots,$$

$$D_j = \begin{vmatrix} a_{11} & a_{12} & 0 & . & . & . & 0 & 0 \\ a_{12} & a_{22} & a_{23} & . & . & . & 0 & 0 \\ . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & . & a_{j,j-1} & a_{jj} \end{vmatrix}.$$

THEOREM II. *A sufficient condition that the bilinear form* (6) *be positive definite is that the sequence* $D_0 = 1$, $D_1$, $D_2,\dots$, $D_n$ *have all of its terms positive. A sufficient condition that the form be negative definite is that the terms of the sequence alternate in sign.*

*Proof.* We note first that

$$D_j = a_{jj}D_{j-1}-a_{j,j-1}^2 D_{j-2}. \tag{7}$$

We next write

$$P(h_1,\dots, h_n) = a_{11}h_1^2+2a_{12}h_1h_2+a_{22}h_2^2+\dots+a_{nn}h_n^2.$$

Assuming that no $D$ is zero, completing the square, proceeding

† See Kowalewski, *Determinantentheorie*, ed. 1909, p. 239, where the general form is discussed.

from the beginning, and using relation (7), we arrive at the following identity:

$$P(h_1,\dots, h_n) = \frac{1}{D_1}(a_{11}^2 h_1^2+2a_{11}a_{12}h_1h_2+a_{12}^2h_2^2)+$$
$$+\frac{D_1}{D_2}\left(\frac{D_2^2}{D_1^2}h_2^2+2\frac{D_2}{D_1}a_{23}h_2h_3+a_{23}^2h_3^2\right)+\dots+$$
$$+\frac{D_{n-2}}{D_{n-1}}\left(\frac{D_{n-1}^2}{D_{n-2}^2}h_{n-1}^2+2\frac{D_{n-1}}{D_{n-2}}a_{n-1,n}h_{n-1}h_n+a_{n-1,n}^2h_n^2\right)+$$
$$+\frac{D_n}{D_{n-1}}h_n^2.$$

The conclusion is drawn immediately from this relation.

*Our condition states that if a solution of the recurrent relation* (7) *subject to the initial conditions*

$$D(0) = 1, \qquad D(1) = a_{11} = \phi_{y(1)y(1)}$$

*is positive throughout the interval over which $y(i)$ is variable then $y(i)$ renders* (2) *a minimum. If, on the other hand, $D_j$ alternates in sign throughout this interval $y(i)$ renders* (2) *a maximum.*

By differentiation of (2) we have

$$a_{ii} = \phi_{y(i)y(i)} = f_{pp}(i)+f_{pp}(i-1)-2f_{yp}(i)+f_{yy}(i), \tag{8}$$

$$a_{i,i-1} = \phi_{y(i)y(i-1)} = f_{yp}(i-1)-f_{pp}(i-1). \tag{9}$$

## 3. A simple example

Consider the sum $$\sum_{i=0}^{n-1}(4y^2+3p^2). \tag{10}$$

Here

$f(i) = 4y^2+3p^2$, $f_y = 8y$, $f_p = 6p$, $f_{yy} = 8$, $f_{pp} = 6$, $f_{yp} = 0$.

Equation (4′) takes the form

$$\Delta(8y-6p)+8y = 0,$$

which can be written

$$3y(i+2)-10y(i+1)+3y(i) = 0. \tag{11}$$

To solve this recurrent relation we form the auxiliary equation

$$3m^2-10m+3 = 0,$$

which yields $m = 3, \frac{1}{3}$. The general solution of (11) is then

$$y(i) = c_1 3^i + c_2 \frac{1}{3^i}. \tag{12}$$

Let us impose the boundary conditions

$$y(0) = A, \qquad y(n) = B.$$

Under these conditions equations (5) do not enter. They are necessary only under the conditions that $y(0)$ and/or $y(n)$ are variable. Solving for $c_1$ and $c_2$ we find

$$c_1 = \frac{A - B3^n}{1 - 3^{2n}}, \qquad c_2 = \frac{3^n B - A3^{2n}}{1 - 3^{2n}}. \tag{13}$$

By (8) and (9) we find

$$\phi_{y(i)y(i)} = 20, \qquad \phi_{y(i)y(i-1)} = -6.$$

Equation (7) takes the form

$$D(j+2) - 20D(j+1) + 36D(j) = 0,$$

from which $$D(j) = C_1 18^j + C_2 2^j.$$

Moreover, $D(0) = 1, \qquad D(1) = \phi_{y(1)y(1)} = 20.$

By means of these initial values we determine $C_1$ and $C_2$ obtaining

$$D(j) = \tfrac{1}{16}(18^{j+1} - 2^{j+1}).$$

This is positive when $j$ is positive and consequently $\phi$ is a minimum when $y$ is given by (12) and (13).

## 4. Minimum surface of revolution

Connect the points $(0, A)$ and $(n, B)$ in the $(i, y)$-plane with a broken straight line, each segment of the broken line extending from a point $\{i, y(i)\}$ to a point $\{(i+1), y(i+1)\}$. The broken line is then revolved about the $i$-axis. We wish to determine the function $y(i)$ which renders the surface generated a minimum. Let $A$ and $B$ be positive.

We are to minimize

$$\sum_{i=0}^{n-1} \{y(i+1) + y(i)\}(1+p^2)^{\frac{1}{2}}. \tag{14}$$

Here $f = (p+2y)(1+p^2)^{\frac{1}{2}}$. Equation (4′) becomes

$$\Delta\frac{1-2py}{(1+p^2)^{\frac{1}{2}}} = -2(1+p^2)^{\frac{1}{2}}, \tag{15}$$

which we can write

$$\frac{1-2p(i+1)y(i+1)}{[1+\{p(i+1)\}^2]^{\frac{1}{2}}} = -\frac{1+2p(i)y(i+1)}{\{1+p^2(i)\}^{\frac{1}{2}}}. \tag{15′}$$

Call the right-hand member of this equation $-2C$ and solve for $p(i+1)$. We get

$$p(i+1) = \frac{y(i+1)\pm C(1+4[\{y(i+1)\}^2-C^2])^{\frac{1}{2}}}{2[\{y(i+1)\}^2-C^2]}$$

provided $\{y(i+1)\}^2-C^2 \neq 0$. Replace $p(i+1)$ by

$$y(i+2)-y(i+1)$$

and we get
$$y(i+2) = \frac{N\pm S}{D},$$

where

$$N = 2\{y(i+1)\}^3+2y(i+1)\{y(i)\}^2+y(i+1),$$

$$S = y(i+1)+2\{y(i+1)\}^3-2y(i+1)\{y(i)\}^2+y(i),$$

$$D = 4y(i)y(i+1)-1.$$

We readily show that the vanishing of $D$ is necessary and sufficient for the vanishing of $[y(i+1)]^2-C^2$ which for the moment we have assumed not to vanish.

Substituting for $N$, $S$, and $D$

$$y(i+2) = \frac{2y(i+1)+4\{y(i+1)\}^3+y(i)}{4y(i)y(i+1)-1} \tag{16}$$

or
$$y(i+2) = y(i). \tag{17}$$

By a rather tedious substitution, we find that $y(i+2)$ as given by (16) identically satisfies (15).

Substituting (17) in (15) yields $1+2p(i)y(i+1) = 0$ which is in general not true. However, if we assume it to be true, (16) reduces to (17). In other words (17) never gives $y(i+2)$ unless it is also given by (16). If $4y(i)y(i+1)-1 = 0$, substitution in (15) yields $1+2p(i)y(i+1) = 0$. This is inconsistent with $4y(i)y(i+1)-1 = 0$ for real values of $y(i)$ and $y(i+1)$. We

conclude that the satisfaction of (16) is necessary as well as sufficient for the satisfaction of (15) by real values. If $4y(i)y(i+1)-1$ should be negative $y$ would necessarily change sign, since in this case from (16), if $y(i)$ and $y(i+1)$ are positive or zero $y(i+2)$ is negative. But a function which generates the surface of revolution which is an absolute minimum cannot be both positive and negative, as can be proved by elementary geometry.

We consequently conclude that *for the minimum surface of revolution in question it is necessary that y satisfy* (16) *and that* $4y(i)y(i+1)-1>0$.

For equation (16) we have the following existence theorem:

*Given any two initial values* $y(0)$ *and* $y(1)$, *it is possible to successively determine* $y(2), y(3),\ldots, y(n)$ *provided that at no point* $4y(i)y(i+1)-1 = 0$, $i = 0, 1,\ldots, n-2$. *If at some such point* $4y(i)y(i+1)-1 = 0$ *there is no solution with the given initial values.*

Since $4y(i)y(i+1)-1 = 0$ is a rational algebraic equation in $y(0)$ and $y(1)$, the initial values for which there is no solution are correspondingly restricted.

We have remarked that all solutions of (16) are solutions of (15) and that (15) has no other solutions. We consequently have an existence theorem for (15).

We next show by actual substitution that

$$y = \frac{1}{2\sinh(1/2a)}\cosh\frac{i-k}{a} \tag{18}$$

is a solution of (15).

$$p = \frac{1}{2\sinh(1/2a)}\left(\cosh\frac{i+1-k}{a}-\cosh\frac{i-k}{a}\right) = \sinh\left(\frac{i-k}{a}+\frac{1}{2a}\right),$$

$$[1+p^2]^{\frac{1}{2}} = \cosh\left(\frac{i-k}{a}+\frac{1}{2a}\right),$$

$$\frac{1-2py}{[1+p^2]^{\frac{1}{2}}} = -\frac{1}{\sinh(1/2a)}\sinh\frac{i-k}{a},$$

$$\Delta\frac{1-2py}{[1+p^2]^{\frac{1}{2}}} = -2\cosh\left(\frac{i-k}{a}+\frac{1}{2a}\right),$$

verifying the solution.

We next ask the question: Does (18) give all solutions of (15) which are positive at all points considered and which are described in our existence theorem above as determined by point-to-point solution of (16)?

Suppose $4AC-1 > 0$ and

$$\begin{aligned} y(0) &= A = \frac{1}{2\sinh(1/2a)}\cosh\frac{k}{a}, \\ y(1) &= C = \frac{1}{2\sinh(1/2a)}\cosh\left(\frac{1}{a}-\frac{k}{a}\right). \end{aligned} \tag{19}$$

We wish to solve these two equations for $a$ and $k$.

$$(C-A) = -\sinh\left(\frac{k}{a}-\frac{1}{2a}\right) = -\sinh\frac{k}{a}\cosh\frac{1}{2a}+\cosh\frac{k}{a}\sinh\frac{1}{2a}. \tag{20}$$

Replace $\cosh(k/a)$ in this equation by $2A\sinh(1/2a)$ and $\sinh(k/a)$ by $\pm[4A^2\sinh^2(1/2a)-1]^{\frac{1}{2}}$. Solving for $\sinh(1/2a)$ we find

$$\sinh\frac{1}{2a} = \pm\left[\frac{(C-A)^2+1}{4AC-1}\right]^{\frac{1}{2}}. \tag{21}$$

We first discard the minus sign before the radical, as we are only interested in solutions which are always positive. We next notice that $4AC-1 > 0$, inasmuch as $4y(i+1)y(i)-1 > 0$ in particular when $i = 0$. Equation (21) consequently determines a positive value for $\sinh(1/2a)$ and consequently a positive value of $a$. To determine $k$ replace $\sinh(1/2a)$ by the above value in (20), $\cosh(1/2a)$ by $[1+\sinh^2(1/2a)]^{\frac{1}{2}}$, $\cosh(k/a)$ by $2A\sinh(1/2a)$, and solve for $\sinh(k/a)$. A unique value is determined for $k$. These values satisfy both equations (19) as again is verified by substitution.

*Formula* (18) *then gives us the general solution of* (15) *which remains positive throughout the interval in question and consequently is a necessary form for the function* $y(i)$ *in order that the surface of revolution be a minimum.*

## EXERCISE

Discuss the maxima and minima of

(i) $\sum_{i=0}^{1000}(y^2+2py+6p^2)$, (ii) $\sum_{i=0}^{n}(1+p^2)^{\frac{1}{2}}$.

IX

# THE GENERAL BOUNDARY PROBLEM

## 1. Compatibility and incompatibility

WE shall consider the recurrent relation of the $n$th order:

$$P(y) \equiv p_0(i)y(i+n)+p_1(i)y(i+n-1)+\ldots+p_n(i)y(i) = p(i), \quad (1)$$

where $p_0(i), p_1(i),\ldots, p_n(i), p(i)$ are defined when $i = a, a+1,\ldots, b-1$ and

$$p_0(i)p_n(i) \neq 0 \quad (2)$$

at any point. Now let

$$A^{(j)}(y) = a_0^{(j)}y(a)+a_1^{(j)}y(a+1)+\ldots+a_{n-1}^{(j)}y(a+n-1),$$
$$B^{(j)}(y) = b_0^{(j)}y(b)+b_1^{(j)}y(b+1)+\ldots+b_{n-1}^{(j)}y(b+n-1),$$
$$j = 1, 2,\ldots, n,$$

where the $a$'s and $b$'s are constants. Let

$$W^{(j)}(y) = A^{(j)}(y)+B^{(j)}(y).$$

The restriction (2) can be replaced in much of the work by $p_0(i) \neq 0$ or $p_n(i) \neq 0$. However, (2) is imposed once and for all for symmetry and to save frequent restatement of restrictions on the coefficients.

We shall consider (1) subject to the conditions

$$W^{(j)}(y) = \gamma_j, \quad j = 1, 2,\ldots, n. \quad (3)$$

The $\gamma$'s are constants.

The system consisting of (1) and (3) will be called *homogeneous* if both

$$p(i) \equiv 0 \quad \text{and} \quad \gamma_1 = \gamma_2 = \ldots = \gamma_n = 0.$$

It will be called *non-homogeneous* in the contrary case. It will be called *semi-homogeneous* in case $p(i) \equiv 0$ and at least one $\gamma$ is not zero. When dealing with the non-homogeneous system we shall speak of the corresponding *reduced* system, meaning the homogeneous system formed by replacing $p(i)$ and $\gamma_1, \gamma_2,\ldots, \gamma_n$ by zero. For ease of reference the homogeneous system is written down

$$P(y) = 0, \quad (4)$$

$$W^{(j)}(y) = 0, \quad j = 1, 2,\ldots, n. \quad (5)$$

Conditions (5) will be assumed at all times to be *linearly independent.*

DEFINITION. *The homogeneous system consisting of* (4) *and* (5) *is called incompatible if* (4) *has no solution other than zero which satisfies* (5). *It is said to have k-fold compatibility in case* (4) *has k and not more than k linearly independent solutions which satisfy* (5).

If (4), (5) has $k$-fold compatibility and if $y_1(i),\ldots, y_k(i)$ are linearly independent solutions of the system (4), (5) then the general solution of (4), (5) is given by

$$y(i) = c_1 y_1(i)+c_2 y_2(i)+\ldots+c_k y_k(i),$$

where $c_1, c_2,\ldots, c_k$ are constants; because a solution of (4) given by this expression satisfies (5) and if there were another not given by this formula there would be $k+1$ linearly independent solutions of (4), (5) and hence $(k+1)$-fold compatibility.

We now state in the form of a theorem the following important result.

THEOREM I. *If* $y_1(i),\ldots, y_n(i)$ *are a fundamental system of solutions of* (4), *a necessary and sufficient condition that the system* (4), (5) *be compatible is that the determinant*

$$D = \begin{vmatrix} W^{(1)}(y_1) & . & . & . & W^{(1)}(y_n) \\ . & . & . & . & . \\ W^{(n)}(y_1) & . & . & . & W^{(n)}(y_n) \end{vmatrix}$$

*vanish. A necessary and sufficient condition that* (4), (5) *have k-fold compatibility is that D be of rank* $n-k$.

This theorem follows from that fact that if we substitute the general solution of (4), namely, $c_1 y_1(i)+c_2 y_2(i)+\ldots+c_n y_n(i)$ in (5), there results a system of linear homogeneous equations for the determination of the $c$'s whose determinant is $D$.

The determinant, $D$, is not always zero. In other words the system (4), (5) is not always compatible. To show this choose $a_0^{(1)} = a_1^{(2)} = \ldots = a_{n-1}^{(n)} = 1$ and all the other $a$'s and all the $b$'s equal to zero. We then have

$$D = y_1(a)y_2(a+1)\ldots y_n(a+n-1);$$

and it is possible to choose $y_1(i)$, $y_2(i)$,..., $y_n(i)$ so that this is different from zero.

THEOREM II. *There always exists a solution of* (4) *which satisfies fewer than n conditions of the form* (5).

This theorem is an immediate consequence of the fact that a system of $n-j$, $j = 1,..., n-1$, homogeneous linear equations in $n$ unknowns always has a solution in which not all the unknowns are zero.

We next prove the following theorem:

THEOREM III. *A necessary and sufficient condition that the non-homogeneous system,* (1), (3), *have one and only one solution is that the reduced system be incompatible.*

We first remark that if the system, (1), (3), is compatible its general solution is formed by adding a particular solution of (1), (3) to the general solution of (4), (5). This general solution of (4), (5) may be identically zero. If we denote the general solution of (1) by

$$y = c_1 y_1 + ... + c_n y_n + Y, \tag{6}$$

where $Y$ is a particular solution of (1), (3), we then must choose $c_1,..., c_n$ so that $c_1 y_1 + ... + c_n y_n$ is a solution of (4), (5); whereupon $y$ will be a solution of (1), (3). Inasmuch as (6) is the general solution of (1) there are no other solutions of the system (1), (3).

To prove the condition necessary: There is one and only one solution to (1), (3). We call this $Y$. Hence we must have $c_1 = ... = c_n = 0$. This means that (4), (5) is incompatible.

To prove the condition sufficient: We need only show that if (4), (5) is incompatible then (1), (3) has a solution. There cannot be two since in (6) we have $c_1 = ... = c_n = 0$. To show that if (4), (5) is compatible then (1), (3) has a solution we first consider the semi-homogeneous case in which $p(i) \equiv 0$. Under this circumstance by Theorem II there exists a solution of equation (4) which satisfies all the reduced conditions except one. This cannot be satisfied by the solution in question since the reduced system is incompatible. We let the unfulfilled condition be the $j$th and call the corresponding solution $y_j$. Let $j$ range from

1 to $n$. Then none of the constants $W^{(j)}(y_j)$ is zero. We now readily verify that

$$\frac{\gamma_1 y_1}{W^{(1)}(y_1)}+\frac{\gamma_2 y_2}{W^{(2)}(y_2)}+\dots+\frac{\gamma_n y_n}{W^{(n)}(y_n)}$$

is a function which satisfies all of conditions (3).

We next denote by $y$ a solution of (1) and by $\bar{y}$ the solution of (4) which satisfies conditions (3). That is, $\bar{y}$ is a function whose existence we have just proved. We now immediately verify that $y+\bar{y}$ is a solution of the system (1), (3) and the proof of our theorem is complete.

## 2. Green's functions

Here, as in the corresponding theory for the linear differential equation, the concept of Green's function arises in the attempt to set up a function which, in the case of an incompatible system, comes as near as possible to satisfying the given conditions.

DEFINITION. *By a Green's function $G(i,\xi)$ of the system* (4), (5) *we understand a function of the integral arguments $i$, $\xi$ defined when $a \leqslant i \leqslant b+n-1$ and $a \leqslant \xi \leqslant b-1$ and such that, when $\xi$ is fixed, $G$, regarded as a function of $i$, satisfies the boundary conditions* (5) *and the equation* (4) *at all points of $a \leqslant i \leqslant b-1$ except for the one value $i = \xi$. Moreover when $G(i,\xi)$ is substituted in* (4) *the left-hand member takes the value* 1 *when $i = \xi$.*

We must first investigate the existence of such functions. As previously let $y_1(i),\dots, y_n(i)$ be a fundamental system of solutions of (4), that is, $n$ linearly independent solutions.

Let

$$\begin{aligned} u_1(i) &= c_1 y_1(i)+\dots+c_n y_n(i),\\ u_2(i) &= d_1 y_1(i)+\dots+d_n y_n(i),\end{aligned} \tag{7}$$

where $c_1,\dots, d_n$ are to be independent of $i$ but will depend upon $\xi$.

Let

$$u(i) = \begin{cases} u_1(i), & i = a, a+1,\dots, \xi,\\ u_2(i), & i = \xi+1, \xi+2,\dots, b-1.\end{cases} \tag{8}$$

The function $u(i)$ satisfies (4) at all integral points of the interval $a \leqslant i \leqslant b-1$ with the possible exception of $\xi, \xi-1,\dots, \xi-n+1$. In order that $u(i)$ satisfy (4) also at the points $\xi-1,\dots, \xi-n+1$,

in so far as these points satisfy the inequality $i \geqslant a$, it is necessary and sufficient that

$$u_2(i) = u_1(i), \quad i = \xi+1,\dots, \xi+n-1. \tag{9}$$

Moreover in order that $u$ satisfy the final requirement for a Green's function it is necessary and sufficient that

$$p_0(\xi)[u_2(\xi+n)-u_1(\xi+n)] = 1. \tag{10}$$

Substitute in (9) and (10) from (7) and let

$$z_j(\xi) = p_0(\xi)[d_j(\xi)-c_j(\xi)], \tag{11}$$

where the dependence of $d_j$ and $c_j$ upon $\xi$ is indicated. We get the equivalent conditions

$$z_1(\xi)y_1(\xi+j)+\dots+z_n(\xi)y_n(\xi+j) = 0, \quad j = 1,\dots, n-1,$$
$$z_1(\xi)y_1(\xi+n)+\dots+z_n(\xi)y_n(\xi+n) = 1.$$

This is a system of linear equations which serve to determine $z_1(\xi),\dots, z_n(\xi)$. Determination is unique since the determinant of the coefficients is the Wronskian of $n$ linearly independent solutions. It is well to note that $z_1(\xi),\dots, z_n(\xi)$ are defined only when $a \leqslant \xi \leqslant b-1$ whereas $y_1(i),\dots, y_n(i)$ are defined when

$$a \leqslant i \leqslant b+n-1.$$

We must now see under what conditions the function $u$ can be made to satisfy the boundary conditions. Substitute $u$ as given by (8) in (5). The following equations result:

$$c_1(\xi)A^{(j)}(y_1)+\dots+c_n(\xi)A^{(j)}(y_n)+d_1(\xi)B^{(j)}(y_1)+\dots+$$
$$+d_n(\xi)B^{(j)}(y_n) = 0, \quad j = 1,\dots, n. \tag{12}$$

Replace the $c$'s by their values as given by (11), and (12) goes into

$$p_0(\xi)[d_1(\xi)W_j(y_1)+\dots+d_n(\xi)W_j(y_n)]$$
$$= z_1(\xi)A^{(j)}(y_1)+\dots+z_n(\xi)A^{(j)}(y_n), \quad j = 1, 2,\dots, n. \tag{13}$$

This is a set of linear equations for the determination of the $d$'s with any set of values for the $z$'s. We remark that $z_1(\xi),\dots, z_n(\xi)$ are linearly independent. If they were not the determinant

$$\Delta = \begin{vmatrix} z_1(\xi-n+1) & . & . & . & z_n(\xi-n+1) \\ . & . & . & . & . \\ z_1(\xi) & . & . & . & z_n(\xi) \end{vmatrix}$$

would vanish for all values of $\xi$ from $a+n-1$ to $b-1$. This is impossible, for multiplying $\Delta$ by

$$\begin{vmatrix} y_1(\xi+1) & . & . & . & . & y_n(\xi+1) \\ . & . & . & . & . & . \\ y_1(\xi+n) & . & . & . & . & y_n(\xi+n) \end{vmatrix}$$

we get, combining rows with rows, a determinant which has zeros everywhere below the primary diagonal and ones at every point of this diagonal and which is therefore not zero.

We prove the following theorem:

THEOREM: *A necessary and sufficient condition that equations* (13) *be consistent for all values of $\xi$ for which $a \leqslant \xi \leqslant b-1$ is that the determinant $D$ do not vanish.*

That the condition is sufficient is immediate from (13). To prove the condition necessary we assume $D = 0$ and show that equations (13) are inconsistent. From the vanishing of $D$ we infer that there exist $n$ constants, $k_1,..., k_n$, not all zero such that

$$k_1 W_1(y_j)+...+k_n W_n(y_j) = 0, \quad j = 1,..., n. \tag{14}$$

We now multiply equations (13) respectively by $k_1,..., k_n$ and add. We obtain

$$0 = z_1 \sum_{j=1}^{n} k_j A^{(j)}(y_1)+...+z_n \sum_{j=1}^{n} k_j A^{(j)}(y_n).$$

This holds for integral values of $\xi$, $a \leqslant \xi \leqslant b-1$. But we have just proved that the $z$'s are linearly independent. Consequently

$$k_1 A^{(1)}(y_j)+...+k_n A^{(n)}(y_j) = 0, \quad j = 1,..., n.$$

Subtract these from (14) and utilize the definition of $W^{(j)}$ and we have

$$k_1 B^{(1)}(y_j)+...+k_n B^{(n)}(y_j) = 0, \quad j = 1, 2,..., n.$$

Now write in the formulae for the $A$'s and $B$'s and we have

$$(k_1 a_1^{(1)}+...+k_n a_n^{(1)})y_j(a)+...+$$
$$+(k_1 a_1^{(n)}+...+k_n a_n^{(n)})y_j(a+n-1) = 0, \tag{15}$$
$$(k_1 b_1^{(1)}+...+k_n b_n^{(1)})y_j(b)+...+$$
$$+(k_1 b_1^{(n)}+...+k_n b_n^{(n)})y_j(b+n-1) = 0, \tag{16}$$
$$j = 1,..., n.$$

We can regard these as equations for the determination of the parentheses; equations (15) for the $a$'s and (16) for the $b$'s. The determinants of the systems are the Wronskians of $y_1,\ldots, y_n$ at $a$ and $b$ respectively and are different from zero since $y_1,\ldots, y_n$ are linearly independent. It results that

$$k_1 a_1^{(j)}+\ldots+k_n a_n^{(j)} = 0, \quad j = 1,\ldots, n,$$

and

$$k_1 b_1^{(j)}+\ldots+k_n b_n^{(j)} = 0, \quad j = 1,\ldots, n.$$

This means that conditions (5) are linearly dependent, which is the desired contradiction.

We now immediately have the following fundamental theorem:

THEOREM: *A necessary and sufficient condition for the existence of a Green's function for the system consisting of* (4) *and* (5), *where* (5) *are assumed linearly independent, is that the system be incompatible. Under which circumstance there is one and only one Green's function*

$$u \equiv G(i, \xi).$$

## 3. An application of Green's function

We shall prove in this section the following theorem:

THEOREM. *If the system* (4), (5) *is incompatible then the system*

$$P(y) = p(i),$$

$$W_j(y) = 0, \quad j = 1,\ldots, n,$$

*has one and only one solution which is given by the following formula:*

$$y(i) = \sum_{\xi=a}^{b-1} G(i, \xi)p(\xi). \tag{17}$$

The proof of this theorem is a matter of verification. Substitute (17) in (1). We have

$$\sum_{j=0}^{n} p_{n-j}(i)y(i+j) = \sum_{j=0}^{n} p_{n-j}(i)\sum_{\xi=a}^{b-1} G(i+j, \xi)p(\xi)$$

$$= \sum_{\xi=a}^{b-1}\left\{\sum_{j=0}^{n} p_{n-j}(i)G(i+j, \xi)p(\xi)\right\}. \tag{18}$$

Each of the sums lying in the braces in (18) is zero except when $i = \xi$, whereupon it is $p(\xi) = p(i)$. But $\xi$ runs through all values from $a$ to $b-1$, consequently for any value of $i$, such that

$a \leqslant i \leqslant (b-1)$, (17) satisfies (1). Now as to boundary conditions (5). Here $G(i, \xi)$ satisfies (5) for any value of $\xi$; that is (17) satisfies (5).

## EXERCISE

Are the following systems compatible or incompatible? If a system is compatible determine the order of compatibility. If a system is incompatible set up the corresponding Green's function. Discuss for all value of $j$.

(a)
$$\Delta^2 y(i-1)+4\left[\sin^2\frac{\pi i}{2(n+1)}\right]y(i) = 0,$$
$$y(0) = y(n+1) = 0.$$

(b)
$$\Delta^2 y(i-1)+4\left[\sin^2\frac{\pi i}{2(n+1)}\right]y(i) = 0,$$
$$y(0) = y(n+1),$$
$$y(1) = y(n+2).$$

(c)
$$\Delta^2 y(i-1)+4\left[\sin^2\frac{\pi i}{2(n+1)}\right]y(i) = 0,$$
$$y(0) = -y(n+1),$$
$$y(1) = -y(n+2).$$

## SUGGESTED REFERENCE

BÔCHER, M.: 'Boundary Problems and Green's Functions for Linear Differential and Difference Equations', *Ann. of Math.*, 2nd series, **13**, 71.

# X

# STURM-LIOUVILLE THEORY

## 1. Fundamental theorems on nodes

CONSIDER the difference equation (recurrent relation)

$$\Delta[K(i,\lambda)\Delta y(i)]-G(i,\lambda)y(i+1) = 0, \tag{1}$$

where $K(i, \lambda)$ and $G(i, \lambda)$ are real and defined as functions of $i$ as follows: $G(i, \lambda)$ over

$$0 \leqslant i \leqslant b-1, \tag{2}$$

and $K(i, \lambda)$ over

$$0 \leqslant i \leqslant b. \tag{3}$$

The operator $\Delta$ applies to the variable $i$ which is limited to integral values. The variable $\lambda$ is a real parameter. The functions $K(i, \lambda)$ and $G(i, \lambda)$ are continuous in $\lambda$ over the interval of their definition, which is not specified, and are such that when $\lambda$ increases $G$ increases and $K$ does not increase. $K(i, \lambda)$ is always positive.

We call equation (1) under these conditions Sturm's normal form. We shall consider only real solutions of (1).

The intervals of definition which we have assumed for $K(i,\lambda)$ and $G(i, \lambda)$, namely (3) and (2), are such that a solution with arbitrarily assigned values at $i = 0$ and $i = 1$ is uniquely determined by (1) for all values of $i$ on the interval

$$0 \leqslant i \leqslant b+1.$$

Before proceeding farther, however, for convenience in our proofs, we extend the definitions of $K(i, \lambda)$ and $G(i, \lambda)$ as follows:

$$K(i, \lambda) = 1, \quad i > b;$$

$$G(i, \lambda) = G(b-1, \lambda), \quad i \geqslant b.$$

Under these circumstances a solution of extended equation (1) is determined by its values at 0 and 1 for all positive integral values of $i$. It will be observed that our extended definition of $K(i, \lambda)$ and $G(i, \lambda)$ preserves the essential behaviour of these functions.

We shall consider solutions of (1) for which $y(0)$ and $y(1)$ are

chosen as *continuous functions* of $\lambda$. It results that $y(i)$ is a continuous function of $\lambda$, $i$ being any positive integer. We exclude the possibility $y(0) = y(1) = 0$ for any value of $\lambda$. The only solution of (1) satisfying these equalities is the trivial one $y(i) \equiv 0$. We also exclude $y(1) = 0$ under any circumstances, inasmuch as this as an initial condition is the same as $y(0) = 0$ with the change of variable $i$ into $i-1$. Moreover, we note that for no value of $\lambda$ will $y(j) = y(j+1) = 0$, $j > 0$. If this were true by (1) we would have $y(j-1) = ... = y(1) = y(0) = 0$. But we have excluded this. Now $K > 0$; and consequently if $j > 0$ and $y(j) = 0$, then $y(j-1)$ and $y(j+1)$ are of opposite signs. We conclude that if through continuous variation of $\lambda$ the number of nodes of the solution on the interval $0 < x < b+1$ changes, it must change by the entrance or exit of nodes over an end-point of the interval. It will develop that in § 2 we shall so restrict $y(0)$ that no node enters or leaves the interval at the point 0. We shall require that either $y(0) = 0$ for all values of $\lambda$ or $y(0) = 0$ for no value of $\lambda$.

Let $\bar{\lambda} = \lambda + \Delta\lambda$, where $\Delta\lambda > 0$ is such that $\bar{\lambda}$ still belongs to the domain of definition of $K$ and $G$. Using this value of $\lambda$ we replace equation (1) by

$$\Delta[K(i, \bar{\lambda})\Delta\bar{y}(i)] - G(i, \bar{\lambda})\bar{y}(i+1) = 0. \tag{4}$$

Let us continue to assume now and henceforth that $y(0)$ and $y(1)$ are continuous functions of $\lambda$.

Let $$V(i) \equiv y(i)K(i, \bar{\lambda})\Delta\bar{y}(i) - \bar{y}(i)K(i, \lambda)\Delta y(i). \tag{5}$$

Suppose that $y(0)$ and $y(1)$ are so chosen as functions of $\lambda$ that $V(0) \geqslant 0$.

From (1) and (4),

$$y(i+1)\Delta[K(i, \bar{\lambda})\Delta\bar{y}(i)] - \bar{y}(i+1)\Delta[K(i, \lambda)\Delta y(i)] - \\ -[G(i, \bar{\lambda}) - G(i, \lambda)]\bar{y}(i+1)y(i+1) = 0.$$

It follows that

$$V(i+1) - V(i) = [K(i, \bar{\lambda}) - K(i, \lambda)]\Delta y(i)\Delta\bar{y}(i) + \\ +[G(i, \bar{\lambda}) - G(i, \lambda)]\bar{y}(i+1)y(i+1). \tag{6}$$

Sum both sides of (6) from 0 to $j-1 \geqslant 0$. We get

$$V(j)-V(0) = \sum_{i=0}^{j-1}[K(i,\bar{\lambda})-K(i,\lambda)]\Delta\bar{y}(i)\Delta y(i)+ \\ +\sum_{i=0}^{j-1}[G(i,\bar{\lambda})-G(i,\lambda)]\bar{y}(i+1)y(i+1). \quad (7)$$

For sufficiently small values of $\Delta\lambda$ the products $\bar{y}(i+1)y(i+1)$ and $\Delta\bar{y}(i)\Delta y(i)$ are both non-negative on account of the continuity of $y(i)$ as a function of $\lambda$. Moreover, $y(1) \neq 0$ by hypothesis. Consequently, $\bar{y}(1)y(1) > 0$. Also

$$K(i,\bar{\lambda})-K(i,\lambda) \geqslant 0 \quad \text{and} \quad G(i,\bar{\lambda})-G(i,\lambda) > 0.$$

It results that the right-hand member of (7) is positive. Hence, since $V(0) \geqslant 0$,

$$V(j) > 0, \quad j > 0. \quad (8)$$

Now let us assume $y(j) \neq 0$. Then when $\Delta\lambda$ is sufficiently small,

$$y(j)\bar{y}(j) > 0. \quad (9)$$

From (8) we get

$$\frac{K(j,\bar{\lambda})\Delta\bar{y}(j)}{\bar{y}(j)} > \frac{K(j,\lambda)\Delta y(j)}{y(j)}. \quad (10)$$

*We have thus proved that the ratio*

$$\frac{K(j,\lambda)\Delta y(j)}{y(j)}$$

*is an increasing function of* $\lambda$, *as long as* $y(j) \neq 0$ and $j \geqslant 1$.

If a node lies between $j$ and $j+1$ then $y(j)$ and $\Delta y(j)$ are of opposite sign. Since $K > 0$ and $K(j,\bar{\lambda}) \geqslant K(j,\lambda)$ we conclude from (10) that $-\Delta y(j)/y(j)$ is a positive decreasing function of $\lambda$. This means that $|y(j+1)|/|y(j)|$ is a decreasing function of $\lambda$. In other words if $j \geqslant 1$ *when a node lies between* $j$ *and* $j+1$ *it moves to the right as* $\lambda$ *increases.*

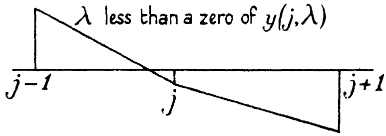
Now let $y(j) = 0$ when $\lambda = \lambda'$, that is, $y(j,\lambda') = 0$. Then, by (8), $\bar{y}(j) \neq 0$ for sufficiently small $\Delta\lambda$. Hence $\lambda'$ is not a cluster point of zeros of $y(j)$ as a function of $\lambda$. Moreover, if $\bar{\lambda} = \lambda'+\Delta\lambda$ with $\Delta\lambda$ sufficiently small, from (8) $\bar{y}(j)K(j,\lambda')\Delta y(j) < 0$. From this, $\bar{y}(j)y(j+1) < 0$. But $y(j+1)\bar{y}(j+1) > 0$. Hence also

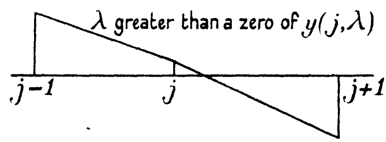
$$\bar{y}(j)\bar{y}(j+1) < 0.$$

In other words, when $\lambda = \lambda' + \Delta\lambda$ the node previously at $j$ is to the right of the point $j$. Also if $\lambda'$ is a zero of $\bar{y}(j)$ as a function of $\lambda$ then

$$y(j)K(j, \lambda'+\Delta\lambda)\Delta\bar{y}(j) > 0,$$

which means that $y(j)\bar{y}(j+1) > 0$ and hence $y(j)y(j+1) > 0$. We conclude that the broken-line graphs of $y(i)$ for values of $\lambda$ close to a zero of $y(j, \lambda)$ are somewhat as follows:




*The node moves to the right through $j$.* When $\lambda$ approaches a zero of $y(j)$ as a function of $\lambda$ from below, the ratio $y(j-1)/y(j)$ becomes negatively infinite, $y(j+1)/y(j)$ and consequently $K(j, \lambda)\Delta y(j)/y(j)$ becomes positively infinite.

If $V(0) > 0$ and $y(0) = 0$ then $\bar{y}(0)\bar{y}(1) < 0$, that is, the node at 0 has advanced to the right. If $V(0) > 0$ and $y(0) \neq 0$, extend (8) and the subsequent reasoning to include $j = 0$. We see that a node on the interval $0 < x < 1$ moves to the right.

We thus have completed the proof of the following important theorem.

THEOREM I. *If $y(i) \not\equiv 0$ is a solution of* (1) *such that $y(0)$ and $y(1)$ are continuous functions of $\lambda$ and such that $V(0)$ remains greater than or equal to zero as $\lambda$ increases, then any node of $y(i)$ on the interval $1 < x < \infty$ moves continuously to the right as $\lambda$ increases.*

*If $V(0)$ remains greater than zero as $\lambda$ increases, any node on the interval $0 \leqslant x \leqslant 1$ moves continuously to the right also.*

## 2. Theorems of oscillation and comparison

Consider two equations,

$$\Delta[K_1(i)\Delta y_1(i)] - G_1(i)y_1(i+1) = 0, \tag{11}$$

$$\Delta[K_2(i)\Delta y_2(i)] - G_2(i)y_2(i+1) = 0, \tag{11'}$$

where $K_2(i) \geqslant K_1(i)$, $G_2(i) > G_1(i)$ are defined over (3) and (2) respectively.

THEOREM II. *If $y_1(i)$ is a real solution of* (11) *and $y_2(i)$ a real solution of* (11′), *neither identically zero, and so chosen that*

$$y_1(0)K_2(0)\Delta y_2(0)-y_2(0)K_1(0)\Delta y_1(0) \geqslant 0, \tag{12}$$

*and either $y_1(0) = y_2(0) = 0$ or $y_1(0).y_2(0) > 0$; then $y_1(i)$ has at least as many nodes on the interval*

$$0 < x \leqslant b+1 \tag{13}$$

*as has $y_2(i)$.*

Let
$$\bar{K}(i, \lambda) = \lambda K_2(i)+(1-\lambda)K_1(i),$$
$$\bar{G}(i, \lambda) = \lambda G_2(i)+(1-\lambda)G_1(i)$$

and consider the equation

$$\Delta[\bar{K}(i, \lambda)\Delta y(i, \lambda)]-\bar{G}(i, \lambda)y(i+1, \lambda) = 0. \tag{11″}$$

We shall assume $y_1(0) = y_2(0)$. There is no loss of generality in this inasmuch as multiplying a solution by a positive constant does not affect nodes nor does it affect (12). Let $y(i, \lambda)$ be a solution of (11″) such that $y(0, \lambda) = y_1(0) = y_2(0)$. Let $y(1, \lambda)$ be a continuous function of $\lambda$ so chosen that $K(0, \lambda)\Delta y(0, \lambda)$ changes monotonically from $K_1(0)\Delta y_1(0)$ to $K_2(0)\Delta y_2(0)$ as $\lambda$ increases from 0 to 1. Now the fact that $y_1(0) = y_2(0) = y(0, \lambda)$ and that $K(0, \lambda)\Delta y(0, \lambda)$ changes monotonically assures $V(0) \geqslant 0$, where $V(i)$ is as defined in (5). Moreover since $y(0, \lambda)$ is constant no node passes through the point 0 as $\lambda$ increases from 0 to 1. Theorem II then follows from Theorem I.

Now let $C$ be a positive constant such that $K(i, \lambda) \geqslant C$, $G(i, \lambda) > C$ for all values of $\lambda$ considered. Consider the equation

$$\Delta[C\Delta y(i)]-Cy(i+1) = 0, \tag{14}$$

which is equivalent to

$$y(i+2)-3y(i+1)+y(i) = 0. \tag{14′}$$

The auxiliary equation to (14′) is

$$\alpha^2-3\alpha+1 = 0$$

and the general solution of (14) is

$$y = C_1\alpha_1^i + C_2\alpha_2^i, \tag{15}$$

where
$$\alpha_1 = \frac{3-\sqrt{5}}{2}, \qquad \alpha_2 = \frac{3+\sqrt{5}}{2}.$$

Both $\alpha_1$ and $\alpha_2$ are positive. Let $y^{(1)}(i)$ be a solution of (14) such that $y^{(1)}(0) = 0$, $y^{(1)}(1) = 1$. The second condition is added for definiteness only, inasmuch as all solutions of (14) which vanish at the origin are proportional. Determine $C_1$ and $C_2$ for $y^{(1)}(i)$. We find

$$y^{(1)}(i) = \frac{\alpha_2^i - \alpha_1^i}{\alpha_2 - \alpha_1}. \tag{16}$$

This is positive for all values of $i > 0$, that is, $y^{(1)}(i)$ has no node on the interval (13). It results from Theorem II that a solution of (1) satisfying the initial condition $y(0) = 0$ has no node on (13).

Now let us go to the other extreme case, still considering initial conditions $y(0) = 0$, $y(1) = 1$. Suppose

$$K(i, \lambda) < C, \qquad G(i, \lambda) < -kC \tag{17}$$

for all $i$. Write down

$$y(i+2) + (k-2)y(i+1) + y(i) = 0. \tag{18}$$

The auxiliary equation to (18) is

$$\alpha^2 + (k-2)\alpha + 1 = 0. \tag{19}$$

Denote its roots by $\alpha_1$ and $\alpha_2$. Let us assume $k$ large and positive. Then, since $-(\alpha_1+\alpha_2) = k-2$ and $\alpha_1\alpha_2 = 1$, one of the numbers is numerically large and the other numerically small. Both are negative. Assume that $\alpha_2$ is the numerically large one. By choosing $k$ sufficiently great the ratio $\alpha_2/\alpha_1$ is made as large as we wish to make it. Now let $y_1^{(2)}(i)$ be the solution of (18) satisfying $y_1^{(2)}(0) = 0$, $y_1^{(2)}(1) = 1$. Then

$$y_1^{(2)}(i) = \frac{\alpha_2^i - \alpha_1^i}{\alpha_2 - \alpha_1}. \tag{20}$$

Choose $k$ so large that the sign of this is determined by $\alpha_2^i$. Then $y_1^{(2)}(i)$ will be positive or negative as $i > 0$ is odd or even. In other words $y_1^{(2)}(i)$ has the maximum possible number of nodes on the interval (13). As a consequence of Theorem II,

if (17) are satisfied, a solution of (1) satisfying the initial condition, $y(0) = 0$, has the maximum possible number of nodes on (13).

We are now in a position to prove the simplest form of the Theorem of Oscillation. We have defined $K(i, \lambda) \equiv 1$ when $i > b$ and $G(i, \lambda) = G(b-1, \lambda)$ when $i > b-1$. Let $y(i)$ be a solution such that $y(0) = 0$, $y(1) = 1$. Assume that as $\lambda$ increases $G(i, \lambda)$ increases from $-\infty$ to a positive value. Then, according to Theorem II and the facts that we have just observed about the equation with constant coefficients, $y(i)$ has first the maximum number of nodes on (13), and finally none at all. Since nodes move continuously to the right we have the following theorem.

THEOREM III. *If $G(i, \lambda)$ increases from $-\infty$ to positive values as $\lambda$ increases there exists one and only one value of $\lambda$, $\lambda_m$, such that a solution of* (1) *vanishing at* 0 *vanishes again at* $b+1$ *with exactly $m$ nodes between* 1 *and $b$. Moreover,*

$$0 \leqslant m \leqslant b-1 \tag{21}$$

*and*

$$\lambda_m > \lambda_{m+1}.$$

The values $\lambda_m$ are called *characteristic values* for equation (1), subject to the boundary conditions $y(0) = y(b+1) = 0$.

We next change the left-hand boundary conditions.

Let us assume the condition

$$[1+\phi(\lambda)]y(0)-K(0, \lambda)\Delta y(0) = 0 \tag{22}$$

with $y(1) \neq 0$.

Assume $1+\phi(\lambda)$ a continuous non-decreasing function, positive† for sufficiently great values of $\lambda$. We denote a particular solution of (1) satisfying (22) by $y(i)$. All such solutions are proportional and hence have the same nodes; but the one that we shall consider satisfies the conditions

$$y(0) = 1, \qquad K(0, \lambda)\Delta y(0) = 1+\phi(\lambda). \tag{23}$$

Consider our comparison equation (18) subject to the initial

† An alternative requirement to $1+\phi(\lambda) > 0$ is $G(i, \lambda) \to \infty$.

conditions $y^{(2)}(0) = 1$, $y^{(2)}(1) = C \equiv 1+\phi(\lambda)$. Denote a solution of (18) satisfying these conditions by $y_2^{(2)}(i)$. Then

$$y_2^{(2)}(i) = \frac{C-\alpha_2}{\alpha_1-\alpha_2}\alpha_1^i + \frac{\alpha_1-C}{\alpha_1-\alpha_2}\alpha_2^i.$$

With $k > 4$ both $\alpha_1$ and $\alpha_2$ are real, negative, and distinct. As previously, if $k$ is sufficiently large $y_2^{(2)}(i)$ will alternate in sign when

$$1 \leqslant i \leqslant b+1. \tag{24}$$

We conclude from Theorem II that if inequalities (17) are satisfied and $k$ is sufficiently large, $y$ has the maximum number of nodes on $1 < x < b+1$.

Now, when $1+\phi(\lambda) > 0$ and $G(i, \lambda) > 0$ at all points of (2), there are no nodes on $1 < x \leqslant b+1$, as is shown by the following reasoning. We rewrite (1) thus:

$$K(i+1, \lambda)\Delta y(i+1) = [K(i, \lambda)+G(i, \lambda)]\Delta y(i)+G(i, \lambda)y(i).$$

Hence $\Delta y(i+1) > 0$ if $y(i) > 0$ and $\Delta y(i) > 0$. But $y(0) > 0$ and $\Delta y(0) > 0$. Hence $\Delta y(1) > 0$. Hence $y(2) > y(1) > 0$. Hence $\Delta y(2) > 0$, etc. Hence $y(i)$ remains of the same sign, $i \geqslant 0$.

Since nodes move continuously to the right as $\lambda$ increases, we are now in a position to state the following theorem.

THEOREM IV. *If $K(i, \lambda) > 0$ does not decrease and $G(i, \lambda)$ increases from $-\infty$ to positive values and $1+\phi(\lambda)$ is a continuous non-decreasing function, and if either $1+\phi(\lambda)$ is positive for $\lambda$ sufficiently great or $G(i, \lambda) \to \infty$ as $\lambda$ increases, or both, then there exists one and only one value of $\lambda$, $\lambda_m$, such that a solution of* (1) *satisfying*

$$[1+\phi(\lambda)]y(0)-K(0, \lambda)\Delta y(0) = 0, \quad y(1) \neq 0,$$

*vanishes at $b+1$ with exactly $m$ nodes on the interval*

$$1 < x < b,$$

*where* $$0 \leqslant m \leqslant b-1 \quad \textit{and} \quad \lambda_m > \lambda_{m+1}.$$

The values $\lambda_m$ are called *characteristic values* for the problem in question. In Theorems III and IV clearly the integer $b+1$ can be replaced by a lesser positive integer such as $b$ with corresponding characteristic values.

We now wish to change the right-hand boundary condition.

Consider (1) subject not only to a left-hand boundary condition

$$[1+\delta\phi(\lambda)]y(0)-\delta K(0,\lambda)\Delta y(0) = 0, \quad y(1) \neq 0, \quad \delta = 0, 1, \tag{25}$$

but also to
$$\psi(\lambda)y(b)-K(b,\lambda)\Delta y(b) = 0, \tag{26}$$

where $\psi(\lambda)$ is a continuous non-increasing function of $\lambda$. Let $y(i)$ be a solution of (1) satisfying (25). We know that

$$K(b,\lambda)\Delta y(b)/y(b)$$

goes from $-\infty$ to $\infty$ as $\lambda$ goes from one zero of $y(b)$, as a function of $\lambda$, to the next. Hence, there is one and only one value of $\lambda$ between each two successive zeros of $y(b)$ for which (26) is satisfied. We denote those values for which $y(b) = 0$, that is, the characteristic values for (1) subject to (25) and $y(b) = 0$, by $\bar\lambda_0, \bar\lambda_1, \ldots, \bar\lambda_{b-2}$. We have thus proved the following theorem:

THEOREM V. *Let $G(i,\lambda)$ increase with $\lambda$ from $-\infty$ to positive values and assume that either $1+\phi(\lambda)$ increases to positive values or $G(i,\lambda) \to \infty$ as $\lambda$ increases, or both; then there exists one and only one value of $\lambda$, which we denote by $\lambda_m^{(1)}$, such that a solution of* (1) *satisfying* (25) *also satisfies* (26) *with exactly $m$ nodes on the interval*

$$1 < x < b, \tag{27}$$

$$1 \leqslant m \leqslant b-2.$$

*Moreover,* $\bar\lambda_0 > \lambda_1^{(1)} > \bar\lambda_1 > \ldots > \lambda_{b-2}^{(1)} > \bar\lambda_{b-2}$.

We now propose the question, Do values $\lambda_\nu^{(1)}$ exist less than $\bar\lambda_{b-2}$ and greater than $\bar\lambda_0$? We put additional restrictions on $G(i,\lambda)$ and in case $\delta \neq 0$ we place an additional restriction on $\phi(\lambda)$ also. These restrictions are: as $\lambda$ increases $1+\phi(\lambda)$ increases from $-\infty$ to positive values, and $G(i,\lambda)$ increases from $-\infty$ to $\infty$ for all $i$.

If $G(i,\lambda)$ is positive and $y(i+1) \geqslant 0$, it follows from (1) that

$$K(i+1,\lambda)\Delta y(i+1)-K(i,\lambda)\Delta y(i) > 0.$$

Since $K(i,\lambda) > 0$, this means that $\Delta y(i+1) > 0$ if $\Delta y(i) \geqslant 0$. Consider first the case that $\delta = 1$, then $y(0) > 0$. Let us assume $1+\phi(\lambda) > 0$ and $y(0) > 0$, then $\Delta y(0) \geqslant 0$ and hence $\Delta y(i) > 0$

for all $i > 0$. Under this circumstance $y(i+1)y(i) > 0$. But, from (1),

$$\frac{K(i+1,\lambda)\Delta y(i+1)}{y(i+1)} = G(i,\lambda)+\frac{K(i,\lambda)\Delta y(i)}{y(i+1)}, \quad y(i+1) \neq 0. \tag{28}$$

But $\dfrac{K(0,\lambda)\Delta y(0)}{y(0)} > 0$. Hence $\dfrac{K(b,\lambda)\Delta y(b)}{y(b)}$ can be made as large as we please by taking $\lambda$ sufficiently large. If $\delta = 0$, that is, $y(0) = 0$, then $y(1) \neq 0$ and when $G$ is positive $\Delta y(i) > 0$ for $i \geqslant 0$. We conclude as before that $\dfrac{K(b,\lambda)\Delta y(b)}{y(b)}$ can be made as large as we please by taking $\lambda$ sufficiently large. In other words, as $\lambda$ increases from $\bar{\lambda}_0$, $\dfrac{K(b,\lambda)\Delta y(b)}{y(b)}$ increases from $-\infty$ to $\infty$. Hence one and only one value $\lambda_0^{(1)} > \bar{\lambda}_0$ exists such that $\dfrac{K(b,\lambda)\Delta y(b)}{y(b)} = \psi(\lambda)$, that is, when $\lambda = \lambda_0^{(1)}$ a solution of (1) satisfying (25) and (26) exists with no nodes on (27).

Now to the other extreme case. Let $\delta = 1$ and $\lambda$ be negative and numerically large. Then $G(i,\lambda)$ is negative and numerically large. If $y(i+1) \neq 0$ we have (28). Moreover, since $\dfrac{K(0,\lambda)\Delta y(0)}{y(0)}$ is numerically large, $\dfrac{K(0,\lambda)\Delta y(0)}{y(1)}$ is numerically small and the sign and order of magnitude of the right-hand member of(28) with $i = 0$ is determined by $G$. By step-by-step procedure we see that, if $\lambda$ is negative and numerically sufficiently large, $\dfrac{K(b,\lambda)\Delta y(b)}{y(b)}$ is negative and numerically as large as we like. In case $\delta = 0$ the same conclusion is also readily drawn. Now inasmuch as $\dfrac{K(b,\lambda)\Delta y(b)}{y(b)}$ becomes positively infinite as $\lambda$ approaches $\bar{\lambda}_{b-2}$ from below,† we conclude that there exists one and only one value $\lambda_{b-1}^{(1)}$ less than $\bar{\lambda}_{b-2}$ such that a solution of (1) satisfying (25) also satisfies (26) with $b-1$ nodes on (27). We consequently have proved the following addition to Theorem V.

† See p. 152.

THEOREM VI. *If to the condition of Theorem V we add:* (1) *for every $i$, $G(i, \lambda)$ increases from $-\infty$ to $\infty$ as $\lambda$ increases;* (2) *if $\delta \neq 0$, $1+\phi(\lambda)$ increases from $-\infty$ to positive values as $\lambda$ increases; then there exists one and only one characteristic value $\lambda_0^{(1)} > \bar{\lambda}_0$ and one and only one characteristic value $\lambda_{b-1}^{(1)} < \bar{\lambda}_{b-2}$.*

## EXERCISE

Sturm's theorem in the theory of algebraic equations is as follows:

Let
$$f(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_n = 0, \tag{1}$$
where $a_1, \dots, a_n$ are real constants, have no multiple roots. Let
$$y(0, \lambda) = f(\lambda), \qquad y(1, \lambda) = f'(\lambda)$$
and form the sequence $y(i, \lambda)$ by means of division as indicated in the following equation:
$$y(i, \lambda) = q(i, \lambda)y(i+1, \lambda) - y(i+2, \lambda), \quad i = 1, 2, \dots, m.$$

Let $V(\lambda)$ be the number of variations in sign (nodes) of $y(i, \lambda)$. Let $c_1$ and $c_2$ ($c_2 > c_1$) be two real numbers neither of which is a root of (1). The number $m$ is chosen not greater than $n$ and such that $y(m, \lambda)$ does not equal zero as $\lambda$ increases from $c_1$ to $c_2$. Then the number of roots of (1) between $c_1$ and $c_2$ is $V(c_1) - V(c_2)$.

Notice that the functions $y(i, \lambda)$, which we can write $y(i)$, satisfy a linear recurrent relation of the second order. Show that when $\lambda$ increases, the number of nodes on the interval $0 \leqslant x \leqslant m$ can change only by exit or entrance of nodes at $x = 0$. Then show that in the neighbourhood of $x = 0$ nodes move continuously to the left as $\lambda$ increases. As a consequence of these facts prove Sturm's theorem.

## SUGGESTED REFERENCES

BÔCHER, M.: *Leçons sur les Méthodes de Sturm*, Paris, Gauthier-Villars, 1917.

—— 'Boundary Problems in One Dimension', Fifth Int. Cong. of Math., *Proceedings*, **1**, 163.

PORTER, M. B.: 'On the Roots of Functions connected by a Linear Recurrent Relation of the Second Order', *Ann. of Math.*, 2nd series, **3**, 55.

# XI

# THE SOLUTION OF A DIFFERENTIAL EQUATION AS THE LIMIT OF THE SOLUTION OF A DIFFERENCE EQUATION

## 1. Introduction

THE early writers on linear differential and difference equations at times established certain results for the difference equation and then, it seems, assumed like results for the differential equation without any careful consideration of the complex limit problem that was involved. The difference equation that they treated was what we know as a recursion formula or recurrent relation, that is, an equation in which the independent variable is limited to integral values or at most a set of isolated congruent values. If the solution of such an equation is plotted in the Cartesian plane, a set of isolated points is the complete graph. If now the difference interval is allowed to approach zero these isolated points can lightly be assumed to approach a continuous curve which is the graph of the solution of the differential equation whose coefficients are the limits of the coefficients of the difference equation. The present chapter treats this limit problem. It is particularly interesting as applying to Sturm-Liouville boundary problems. For compactness in writing we treat only the pair of linear equations of the first order involving two dependent variables. This includes the single equation of the second order. The reader can readily pass to the $n$th order equation if he so desires.

## 2. Fundamental theorem

Consider side by side the sets of equations (1) and (2) below.

$$\begin{aligned} \frac{\Delta U(x)}{h} &= A_{11}(\rho x,\, \rho)U(x)+A_{12}(\rho x,\, \rho)V(x)+B_1(\rho x,\, \rho), \\ \frac{\Delta V(x)}{h} &= A_{21}(\rho x,\, \rho)U(x)+A_{22}(\rho x,\, \rho)V(x)+B_2(\rho x,\, \rho). \end{aligned} \tag{1}$$

$$\begin{aligned} \frac{du}{dx} &= a_{11}(x)u+a_{12}(x)v+b_1(x), \\ \frac{dv}{dx} &= a_{21}(x)u+a_{22}(x)v+b_2(x). \end{aligned} \tag{2}$$

Here $$\Delta\phi(x) = \phi(x+h)-\phi(x).$$

Hypotheses on the coefficients are: (i) $A_{11}(\rho x, \rho),..., B_2(\rho x, \rho)$ are defined when $0 \leqslant x \leqslant 1$ and $\rho > 0$; (ii) $a_{11}(x),..., b_2(x)$ are defined when $0 \leqslant x \leqslant 1$ and $|a_{11}(x)| < a,..., |b_2(x)| < a$; (iii) $|a_{11}(x)|,..., |b_2(x)|$ are Riemann integrable from 0 to 1; (iv) when $\rho \to \infty$, $A_{11}(\rho x, \rho) \to a_{11}(x),..., B_2(\rho x, \rho) \to b_2(x)$, all uniformly in $x$ over the interval $0 \leqslant x \leqslant 1$.

There is, of course, much arbitrariness in the interval for $x$. It can be changed by a linear transformation if desired. The assumption $\rho > 0$ might have been written $\rho > M > 0$. The somewhat peculiar form in which the coefficients in (1) are written, namely $(\rho x, \rho)$, is chosen for convenience and is in no way more restrictive than simply $(x, \rho)$. Thus

$$\phi(x, \rho) = \phi\left(\frac{\rho x}{\rho}, \rho\right) = \Phi(\rho x, \rho).$$

We wish to consider (1) subject to initial conditions

$$U(0) = \tau(\rho), \qquad V(0) = \sigma(\rho)$$

and (2) subject to $u(0) = \tau$, $v(0) = \sigma$. It will appear in the sequel that $\rho = 1/h$.

A special case of (1), (2) is, of course,

$$\frac{\Delta^2 Y(x)}{h^2}+P(\rho x, \rho)\frac{\Delta Y(x)}{h}+Q(\rho x, \rho)Y(x+h) = R(\rho x, \rho), \quad (3)$$

$$\frac{d^2y}{dx^2}+p(x)\frac{dy}{dx}+q(x)y = r(x). \quad (4)$$

The well-known process of successive approximations is now applied to (2) and to (1), as is frequently done in establishing the existence of solutions of (2). For ease in reading the work is carried through first for (2) and then in less detail for (1). This is done rather than treat both sets of equations at once by means of a general operator. Assume initial approximations $u_1(x) = \tau$ and $v_1(x) = \sigma$. Suppose $|\tau| \leqslant c$, $|\sigma| \leqslant c$. Substitute in (2) as follows:

$$\frac{du_2}{dx} = a_{11}(x)u_1(x)+a_{12}(x)v_1(x)+b_1(x),$$

$$\frac{dv_2}{dx} = a_{21}(x)u_1(x)+a_{22}(x)v_1(x)+b_2(x).$$

Here $u_2$ and $v_2$ are second approximations. Solve for $u_2$ and $v_2$ subject to initial conditions $u_2(0) = \tau$, $v_2(0) = \sigma$. Substitute in the right-hand members of (2), denoting the $u$ and $v$ in the left-hand members by $u_3$ and $v_3$. Continue. As is well known, $u_n$ and $v_n$ converge uniformly in $x$ to functions which satisfy the differential equations (2) and the initial conditions. The work is carried out by forming the infinite series

$$u_1+(u_2-u_1)+(u_3-u_2)+\dots$$

and

$$v_1+(v_2-v_1)+(v_3-v_2)+\dots,$$

which we write

$$u_1^{(1)}+u_2^{(1)}+u_3^{(1)}+\dots \tag{5}$$

and

$$v_1^{(1)}+v_2^{(1)}+v_3^{(1)}+\dots. \tag{6}$$

Then $u_n^{(1)}$ and $v_n^{(1)}$, $n > 1$, satisfy the differential equations

$$\frac{du_n^{(1)}}{dx} = a_{11}(x)u_{n-1}^{(1)}+a_{12}(x)v_{n-1}^{(1)},$$

$$\frac{dv_n^{(1)}}{dx} = a_{21}(x)u_{n-1}^{(1)}+a_{22}(x)v_{n-1}^{(1)},$$

and the initial conditions $u_n^{(1)}(0) = 0$, $v_n^{(1)}(0) = 0$. Then

$$|u_2^{(1)}| \leqslant 2ac\int\limits_0^x dx = 2acx \leqslant 2ac,$$

$$|v_2^{(1)}| \leqslant 2ac\int\limits_0^x dx = 2acx \leqslant 2ac.$$

In general

$$|u_{j+1}^{(1)}| \leqslant \int\limits_0^x \frac{c(2a)^j}{(j-1)!}x^{j-1}\,dx = \frac{c(2ax)^j}{j!} \leqslant \frac{c(2a)^j}{j!}, \tag{7}$$

$$|v_{j+1}^{(1)}| \leqslant \frac{c(2a)^j}{j!}, \quad j \geqslant 1,\ 0 \leqslant x \leqslant 1. \tag{8}$$

We know that $c(2a)^j/j!$ is the general term of a convergent series of positive constants. Consequently, by the familiar Weierstrass test, series (5) converges uniformly in $x$ to a function which we denote by $u(x)$ and (6) to a function which we denote by $v(x)$:

$$\lim_{n\to\infty} u_n(x) = u(x),$$

$$\lim_{n\to\infty} v_n(x) = v(x).$$

Now $u(x)$ and $v(x)$ satisfy the differential equations and the initial conditions.

We now apply a like process of successive approximations to (1). To do this we replace $x$ by $i/\rho = hi$, where $i$ is a positive integer or zero. This limits $x$ to the set of congruent values, $0, h, 2h, \ldots$. Replace $U(hi)$ by $w(i)$ and $V(hi)$ by $z(i)$. For the time being hold $\rho = 1/h$ fixed. We have

$$\frac{\Delta w(i)}{h} = A_{11}(i,\rho)w(i)+A_{12}(i,\rho)z(i)+B_1(i,\rho),$$

$$\frac{\Delta z(i)}{h} = A_{21}(i,\rho)w(i)+A_{22}(i,\rho)z(i)+B_2(i,\rho).$$

Here with independent variable $i$ the difference interval is 1.

Proceed as previously, forming successive approximations from initial approximations,

$$w_1(i) \equiv \tau(\rho), \qquad z_1(i) \equiv \sigma(\rho),$$

where $$\tau(\rho) \to \tau, \qquad \sigma(\rho) \to \sigma,$$

when $\rho \to \infty$, and $|\tau(\rho)| \leqslant c$, $|\sigma(\rho)| \leqslant c$. Note that

$$0 \leqslant hi = x \leqslant 1.$$

Assume that

$$|A_{11}(i,\rho)| < a, \quad \ldots, \quad |B_2(i,\rho)| < a.$$

Form the series

$$w_1^{(1)}(i)+w_2^{(1)}(i)+w_3^{(1)}(i)+\ldots, \tag{9}$$

$$z_1^{(1)}(i)+z_2^{(1)}(i)+z_3^{(1)}(i)+\ldots. \tag{10}$$

We know that if $\Delta f(i) = \phi(i)$ and $f(0) = 0$, then

$$f(i) = \sum_{i=0}^{i-1} \phi(i), \quad i > 0.$$

We consequently have

$$|w_{j+1}^{(1)}(i)| = \left|\sum_{i=0}^{i-1} \{A_{11}(i,\rho)w_j^{(1)}(i)+A_{12}(i,\rho)z_j^{(1)}(i)\}h\right|$$

$$\leqslant \sum_{i=0}^{i-1} \frac{h^j c(2a)^j i^{j-1}}{(j-1)!} \leqslant \frac{h^j c(2a)^j i^j}{j!} \leqslant \frac{c(2a)^j}{j!}, \tag{11}$$

$$|z_{j+1}^{(1)}(i)| \leqslant \frac{c(2a)^j}{j!}, \quad j \geqslant 1,\ i \geqslant 0. \tag{12}$$

The right-hand members in (11) and (12) are the same as the right-hand members in (7) and (8). Now replace $i$ by $\rho x$.

Formulae for $w_j^{(1)}(\rho x)$ and $z_j^{(1)}(\rho x)$ are not written out. However, consider the forms which would result from the method of defining these functions. The formula for $w_j^{(1)}$ would be exactly the same as the formula for $u_j^{(1)}$ except that there is an interchange of letters and the symbol $\int_0^x dx$ is replaced by $\sum_{i=0}^{i-1} h$. Now if $\rho \to \infty$ (i.e. $h \to 0$) independently of $x$ the multiple sums appearing in the formula for $w_j^{(1)}$ will each approach a multiple integral which is equal to the corresponding iterated integral appearing in the formula for $u_j^{(1)}$. We assume that the coefficients appearing in (2) are such as to assure this. Continuity is sufficient. Now series (5) and (9) considered as one series in $\rho$ and $x$ converges uniformly by the Weierstrass test. Consequently the series of the limits is equal to the limit of the series:

$$\lim_{\rho\to\infty} w(\rho x) = \lim_{h\to 0} U(x) = u(x).$$

Similarly $$\lim_{\rho\to\infty} z(\rho x) = \lim_{h\to 0} V(x) = v(x).$$

We consequently conclude that the following theorem is true.

Theorem I. *When $\rho \to \infty$ the solution $(U, V)$ of (1) satisfying initial conditions $U(0) = \tau(\rho)$, $V(0) = \sigma(\rho)$ approaches the solution $(u, v)$ of (2) satisfying initial conditions $u(0) = \tau$, $v(0) = \sigma$. The approach is uniform in $x$ when $0 \leqslant x \leqslant 1$.*

We note that $\tau(\rho) \to \tau, \quad \sigma(\rho) \to \sigma.$

Theorem I implies as a special case the following theorem.

Theorem II. *The solution $Y(x)$ of (3) satisfying initial conditions, $Y(0) = \tau(\rho)$, $\rho\Delta Y(0) = \sigma(\rho)$ approaches uniformly in $x$ over the interval $0 \leqslant x \leqslant 1$ the solution $y(x)$ of (4) satisfying initial conditions, $y(0) = \tau$, $\left.\frac{dy}{dx}\right]_{x=0} = \sigma$; and the first difference quotient $\frac{1}{h}\Delta Y(x)$ approaches $\frac{dy}{dx}$ uniformly.*

## 3. Sturm's normal form

Consider

$$\frac{1}{h^2}\Delta[k(\rho x, \rho, \lambda)\Delta y(x)] - g(\rho x, \rho, \lambda)y(x+h) = 0 \tag{13}$$

side by side with

$$\frac{d}{dx}\left[K(x,\lambda)\frac{dy}{dx}\right]-G(x,\lambda)y(x) = 0. \tag{14}$$

Assume that when $\rho \to \infty$,

$$k(\rho x, \rho, \lambda) \to K(x,\lambda) > 0 \quad \text{and} \quad g(\rho x, \rho, \lambda) \to G(x,\lambda),$$

both uniformly, $0 \leqslant x \leqslant 1$. Conditions on (3) are fulfilled and a solution of (13) satisfying initial conditions of the type discussed in § 2 and its first difference quotient approach uniformly respectively the corresponding solution of (14) and its first derivative. If now $\rho$ and $\lambda$ are so related that for the various values of $\rho$ the solution in question is a particular characteristic function of a Sturm-Liouville difference problem with boundary conditions at 0 and 1, the function which it approaches is the corresponding characteristic function of the limiting Sturm-Liouville differential problem.

We give in detail one theorem.

The following theorem is quoted from Chapter X.

Theorem III. *Given*

$$\Delta[k(x,\lambda)\Delta y(x)]-g(x,\lambda)y(x+h) = 0, \tag{15}$$

*where* $x = 0, h, 2h,..., nh = 1$. *If* $k(x,\lambda) > 0$ *does not decrease and* $g(x,\lambda)$ *increases from* $-\infty$ *to positive values as* $\lambda$ *increases, both being continuous in* $\lambda$, *there exists one and only one value of* $\lambda$ *denoted by* $\lambda_m$ *such that a solution of* (15), *not identically zero, vanishing at* 0 *vanishes again at* 1 *with exactly* $m$ *nodes between* 0 *and* 1. *Moreover* $0 \leqslant m \leqslant n-2$ *and* $\lambda_{m-1} > \lambda_m$.

By a node is meant a point where the broken line formed by connecting in order by straight-line segments the points $y(0), y(h), y(2h),..., y(1)$, when plotted in the Cartesian plane, crosses the axis. Now, inasmuch as zeros of a solution of (14) which is not identically zero cannot cluster, we immediately infer the existence of a solution of (14), zero at the origin and at 1, with exactly $j$ zeros on the interval $0 < x < 1$ and corresponding characteristic value $\bar{\lambda}_j$. Since $n \to \infty$ as $h \to 0$ there are an infinite number of such values. These are all distinct. Because with fixed $\lambda$ all solutions of (14) for which $y(0) = 0$

are proportional, and consequently it is impossible for two such solutions to exist, the one with $j$ zeros and the other with $j+l$, $l > 0$. Moreover, the characteristic values must be in the same order as the characteristic values for the difference system, namely $\bar{\lambda}_{j-1} > \bar{\lambda}_j$.

Stated as a theorem we have:

THEOREM IV. *Given equation* (14), *where* $K(x, \lambda) > 0$ *does not decrease as* $\lambda$ *increases and* $G(x, \lambda)$ *actually increases from* $-\infty$ *to positive values, then there exists one and only one real number* $\lambda_m$ *such that when* $\lambda = \lambda_m$ *all solutions not identically zero vanishing at* 0 *vanish again at* 1 *with exactly* $m$ *zeros on the interval*

$$0 < x < 1, \qquad m = 0, 1, 2, \ldots,$$

*and* $\lambda_{m-1} > \lambda_m$.

Let

$$k(\rho x, \rho, \lambda) = K\left(\frac{\rho x}{\rho}, \lambda\right),$$

$$g(\rho x, \rho, \lambda) = G\left(\frac{\rho x}{\rho}, \lambda\right),$$

and carry through in detail the reasoning just given.

## EXERCISES

1. State and prove a generalization of Theorem I.

2. State and prove Sturm-Liouville theorems for the differential equation other than the theorem given in the text.

## SUGGESTED REFERENCE

PICARD, E.: *Traité d'Analyse*, **3**, 90.

# XII

# THE WEIGHTED VIBRATING STRING AND ITS LIMIT

THE problem of the weighted vibrating string goes back at least to Johann Bernoulli. Other references are given at the end of the chapter. The string with fixed end-points is treated as seeming of greatest interest and the special case of equally spaced particles is treated in detail. It is evident that the vibrating weighted string treated in § 1 could be subjected to other boundary conditions than the simple ones $u(0) = u(n+1) = 0$. No change is thereby introduced into our general discussion other than a change in the Sturm-Liouville problem. Also, passage to the limit, in many instances at least, involves but few additional difficulties.

## 1. The general case

Consider an elastic string of negligible mass with fixed end-points under tension and loaded with $n$ particles, each of mass $m$, spaced at irregular intervals and vibrating in a plane. Denote the tension in the string between the $(i-1)$th and the $i$th particles by $T_i$, the coordinates of the $i$th particle by $(x_i, u_i)$ and the length of string between the $i$th and $(i+1)$th particles by $d_i$.

Neglecting gravity, the equations of motion are

$$m\frac{d^2u_i}{dt^2} = -T_i\frac{u_i-u_{i-1}}{d_{i-1}}-T_{i+1}\frac{u_i-u_{i+1}}{d_i}, \tag{1}$$

$$m\frac{d^2x_i}{dt^2} = -T_i\frac{x_i-x_{i-1}}{d_{i-1}}+T_{i+1}\frac{x_{i+1}-x_i}{d_i}. \tag{2}$$

We add the terminal conditions

$$u_0 = u_{n+1} = 0. \tag{3}$$

Let us assume the displacements so small that each particle remains essentially in the same vertical line. Let $\theta_i$ be the acute angle made by the $i$th section of the string with the horizontal. And, under the assumption of the small displacement, let us

replace $\cos\theta_i$ by 1 and $\sin\theta_i$ by $\tan\theta_i$. We then have from (2) that $T_i = T_{i-1} = T_1$. We assume $T_i$ independent of the time and replace it by $T$. Equations (1) and (3) reduce to

$$\frac{d^2u_i}{dt^2} = k(i-1)u_{i-1}-\{k(i-1)+k(i)\}u_i+k(i)u_{i+1}, \tag{4}$$

$$u_0 = u_{n+1} = 0, \tag{5}$$

where $$k(j) = \frac{T}{m(x_{j+1}-x_j)}, \quad j = 1,\ldots, n.$$

In order to solve the system of equations consisting of (4) and (5), we let $u_i = y(i)e^{\mu t}$, where $y(i)$ is independent of $t$. This yields

$$k(i-1)y(i-1)-\{k(i-1)+k(i)+\mu^2\}y(i)+k(i)y(i+1) = 0,$$
$$i = 1,\ldots, n, \tag{6}$$

$$y(0) = y(n+1) = 0. \tag{7}$$

The satisfaction of (6) and (7) is necessary and sufficient that a set of functions of the type $u_i = y(i)e^{\mu t}$ exist satisfying (4) and (5).

Equation (6) is a recurrent relation (difference equation) of the type which is called Sturm's normal form. It is to be solved subject to the boundary conditions (7). Equation (6) can be written

$$\Delta[k(i-1)\Delta y(i-1)]-\mu^2 y(i) = 0. \tag{6'}$$

There exist values of $\mu^2$,

$$\mu_1^2 > \mu_2^2 > \ldots > \mu_n^2,$$

for which solutions of (6) and (7), not identically zero, exist and a complete theorem of oscillation is known.† Each of these values of $\mu^2$ is negative. To prove this, assume the contrary, namely that (6) and (7) are satisfied and $\mu^2 \geqslant 0$. All solutions of (6) satisfying $y(0) = 0$ are proportional. Consider a $y$ such that $y(0) = 0$ and $y(1) = 1$. Then $k(1)y(2) = k(0)+k(1)+\mu^2$. Remark that $k(i) > 0$. Then $y(2) > y(1) > 0$. Then, from (6),

$$y(3) = y(2)+\frac{1}{k(2)}[k(1)\{y(2)-y(1)\}+\mu^2 y(2)]$$

† See Chapter X.

and hence $y(3) > y(2)$. In precisely similar manner

$$y(4) > y(3) > 0, \quad \text{etc.}$$

But this is inconsistent with $y(n+1) = 0$. This is the desired contradiction.

If $\mu^2 = \mu_j^2$, replace $\mu_j^2$ by $-\sigma_j^2$ where $\sigma_j$ is real. Represent the corresponding solution of (6) by $y_j(i)$.

The general solution of the system of differential equations (4) is then

$$u_i = \sum_{j=1}^{n} y_j(i)[c_j^{(1)} \cos \sigma_j t + c_j^{(2)} \sin \sigma_j t]. \tag{8}$$

We note that the summand is $y_j(i)$ multiplied by a function of $t$. In other words, the shape of the string at any time $t$ is determined by the difference equation.

The problem of determining the $c_j$'s from specific initial conditions next presents itself. Let us assume that when $t = 0$, $u_i = U_n(i)$ and $du_i/dt = V_n(i)$. Then

$$\sum_{j=1}^{n} c_j^{(1)} y_j(i) = U_n(i),$$

$$\sum_{j=1}^{n} c_j^{(2)} \sigma_j y_j(i) = V_n(i).$$

Now the functions $y(i)$ constitute an orthogonal set in the sense that

$$\sum_{i=1}^{n} y_j(i) y_p(i) = 0, \quad j \neq p.$$

This is readily proved as follows. Write (6′) for $\mu_j$ and $\mu_p$; multiply respectively by $y_p(i)$ and $y_j(i)$ and subtract; we get

$$\Delta[k(i-1)\{y_p(i-1)\Delta y_j(i-1) - y_j(i-1)\Delta y_p(i-1)\}]$$
$$= (\mu_p^2 - \mu_j^2) y_j(i) y_p(i).$$

Sum this from $i = 1$ to $n$ and note that

$$y_j(0) = y_p(0) = y_j(n+1) = y_p(n+1) = 0$$

and we have the desired result.

Hence

$$\left.\begin{aligned} c_j^{(1)} &= \frac{1}{\sum_{i=1}^{n} \{y_j(i)\}^2} \sum_{i=0}^{n} U_n(i) y_j(i), \\ c_j^{(2)} &= \frac{1}{\sigma_j \sum_{i=1}^{n} \{y_j(i)\}^2} \sum_{i=0}^{n} V_n(i) y_j(i). \end{aligned}\right\} \tag{9}$$

Moreover, the functions $y_j(i)$ can be normalized if we like, that is, each can be multiplied by such a constant that

$$\sum_{i=1}^{n} \{y_j(i)\}^2 = 1.$$

## 2. Special case

An interesting special case is that in which all the particles are equally spaced as among each other and from the fixed end-points of the string. In this case all the $k$'s are equal. Let $M$ be the total mass of all $n$ particles; let the horizontal distance between two successive particles be $\omega$. The horizontal distance from one fixed end-point to the most remote particle is $s$. Then

$$k(i) = \frac{T}{m\omega} = \frac{Tn^2}{Ms} = n^2R^2.$$

Equation (6) and boundary conditions (7) reduce to

$$\Delta^2 y(i-1)+\frac{\sigma^2}{n^2R^2}y(i) = 0, \tag{10}$$

$$y(0) = y(n+1) = 0, \tag{11}$$

where $$\sigma^2 = -\mu^2, \qquad R^2 = \frac{T}{Ms}.$$

If $$\sigma^2 = 4n^2R^2\sin^2\frac{\pi j}{2(n+1)}$$

one readily verifies that a solution of (10) satisfying (11) is

$$y(i) = \sin\frac{i\pi j}{n+1}, \quad j = 1,..., n.$$

This is $y_j(i)$ of the general discussion. Formulae (8) now become

$$u_i = \sum_{j=1}^{n}\left(\sin\frac{i\pi j}{n+1}\right)\left[c_j^{(1)}\cos 2nR\left\{\sin\frac{\pi j}{2(n+1)}\right\}t+\right.$$
$$\left.+c_j^{(2)}\sin 2nR\left\{\sin\frac{\pi j}{2(n+1)}\right\}t\right]. \tag{12}$$

Formulae for the coefficients are as follows:

$$\begin{aligned} c_j^{(1)} &= \frac{2}{n+1}\sum_{i=1}^{n} U_n(i)\sin\frac{i\pi j}{n+1}, \\ c_j^{(2)} &= \frac{1}{2nR\sin\dfrac{\pi j}{2(n+1)}}\,\frac{2}{n+1}\sum_{i=1}^{n} V_n(i)\sin\frac{i\pi j}{n+1}. \end{aligned} \tag{13}$$

We may, if we like, consider the particles as equally spaced within the interval (0, 1) and replace $i/(n+1)$ by $x$. We then allow $n$ to become infinite, retaining a fixed total mass $M$. We have, assuming that (12), thought of as an infinite series, converges uniformly† in $n$,

$$u = \sum_{j=1}^{\infty} (\sin j\pi x)[\bar{c}_j^{(1)} \cos j\pi Rt + \bar{c}_j^{(2)} \sin j\pi Rt], \quad R^2 = T/M. \tag{14}$$

Here

$$\bar{c}_j^{(1)} = 2 \int_0^1 U(x)(\sin j\pi x)\, dx,$$

$$\bar{c}_j^{(2)} = \frac{2}{R\pi j} \int_0^1 V(x)(\sin j\pi x)\, dx,$$

$$U(x) = \lim_{n\to\infty} U_n\{(n+1)x\}, \qquad V(x) = \lim_{n\to\infty} V_n\{(n+1)x\},$$

which last two limits are assumed uniform in $x$.

## 3. The limit in the general case

In equation (6′) replace $\mu^2$ by $-\sigma^2$, then consider the general system

$$\Delta[k(i-1)\Delta y(i-1)] + \sigma^2 y(i) = 0, \tag{15}$$

subject to $$y(0) = y(n+1) = 0.$$

Here $$k(i) = \frac{T}{m(x_{i+1}-x_i)}.$$

Let all particles be on the interval (0, 1). That is, let $s = 1$. Let $n \to \infty$ and the particles approach coincidence but with irregular spacing, the limit being a continuous string with ends at 0 and 1. Let $M = mn$ be constant. Now consider (15) and make the transformation $i = nx$ and let $n \to \infty$. Assume that $\lim_{n\to\infty} n^{-2}k(nx)$ exists and equals $K(x)$. The existence of this limit and an interpretation of $K$ will be discussed presently.

† Summation by parts will show that $c_j^{(1)}$ and $c_j^{(2)}$ are for many functions of the order of $1/n^2$. Integration by parts will show that the same thing is true for $\bar{c}^{(1)}$ and $\bar{c}^{(2)}$.

We have the limiting equation

$$\frac{d}{dx}\left\{K(x)\frac{d\bar{y}}{dx}\right\}+\rho^2\bar{y} = 0, \tag{16}$$

subject to $\bar{y}(0) = \bar{y}(1) = 0$. This is precisely the problem discussed in Chapter XI.

It is well to remark that (16) is not the equation of the string at any time. It is the limiting form of a recurrent relation (difference equation) with independent variable ranging over isolated equally spaced points.

We wish to show the existence of $K(x)$ and at the same time to get an interpretation of it in terms of the physical properties of the string. Let the abscissa of a point on the limit string measured from one end be $q$. Denote the mass of the string measured from the end where $q = 0$ by $\mathfrak{M}(q)$. Let the density of the string be $\delta(q)$, which we assume to be integrable. We have denoted the abscissa of the $i$th particle by $x_i$. Let $x_i$ be given. This determines $i$ for a given $n$ (number of particles). Let $x_i \to q$. Suppose $x_i = F(i) = F(nx) \to \phi(x) = q$. Then

$$n(x_i-x_{i-1}) = \frac{F(nx)-F\{n(x-1/n)\}}{1/n} \to \frac{dq}{dx}.$$

We assume $F$ such a function as to justify this.

But $$n(x_i-x_{i-1}) \to \frac{M}{\delta(q)}.$$

Hence $$\frac{dq}{dx} = \frac{M}{\delta(q)}.$$

Then $$\mathfrak{M}(q) \doteq \int\limits_0^q \delta(q)\,dq = Mx.$$

Then $$\frac{dq}{dx} = \frac{d}{dx}\mathfrak{M}^{-1}(Mx)$$

and $$K(x) = \frac{T}{M(d/dx)\mathfrak{M}^{-1}(Mx)}.$$

Also, since

$$n(x_i-x_{i-1}) \to \frac{M}{\delta(q)}, \qquad K(x) = T\delta(q)M^{-2}.$$

We assume $0 < \epsilon < \delta(q) < A$, and consequently

$$0 < \bar{\epsilon} < K(x) < \bar{A}.$$

Denote the characteristic values for the differential system by $\rho_j$, $j = 1, 2,\ldots$. It is well known† that

$$0 < c_1 j \leqslant \rho_j \leqslant c_2 j, \tag{17}$$

where $c_1$ and $c_2$ are independent of $j$. Denote the characteristic values for the difference problem by $\sigma_j$, $j = 1, 2,\ldots, n$. A relation similar to (17) holds for $\sigma_j$, namely

$$0 < c_1 j \leqslant \sigma_j \leqslant c_2 j, \tag{18}$$

where $c_1$ and $c_2$ are independent of $j$ and $n$. To show this consider the equations

$$\Delta[\bar{\epsilon}\Delta y(i-1)]+\sigma^2 y(i) = 0 \tag{19}$$

and

$$\Delta[\bar{A}\Delta y(i-1)]+\sigma^2 y(i) = 0 \tag{20}$$

side by side with (15), all subject to the boundary conditions $y(0) = y(n) = 0$. Denote the characteristic values for (19) by $\lambda_j$, $j = 1, 2,\ldots$, and for (20) by $\nu_j$, $j = 1, 2,\ldots$. Then

$$\lambda_j > \sigma_j > \nu_j.$$

But the characteristic values for (19) and (20), which are equations with constant coefficients, have been found in § 2 of the present chapter. The desired result is immediate from these values.

We now consider $c_j^{(1)}$ and $c_j^{(2)}$ as given by (9). Perform the transformation $i = nx$. Initial functions are $U_n(nx)$ and $V_n(nx)$. We prove the following theorem:

THEOREM I. *Assume that when* $n \to \infty$, $U(nx) \to U(x)$,

$$(n+1)\Delta U_n(nx) \to \frac{dU(x)}{dx}, \qquad (n+1)^2\Delta^2 U_n(nx) \to \frac{d^2U(x)}{dx^2},$$

$$V_n(nx) \to V(x), \quad \textit{and} \quad (n+1)\Delta V_n(nx) \to \frac{dV(x)}{dx},$$

† See, for example, E. W. Hobson, *Proceedings of London Mathematical Society*, series 2, 6, 378.

*all uniformly in x. Assume moreover that* $U(x)$, $\dfrac{dU(x)}{dx}$, $\dfrac{d^2U(x)}{dx^2}$, $V(x)$, *and* $\dfrac{dV(x)}{dx}$ *are all bounded* $0 \leqslant x \leqslant 1$. *Then*

$$|c_j^{(p)}| \leqslant \frac{B}{j^2}, \quad p = 1, 2, \tag{21}$$

*where B is independent of j and n.*

This result follows from applying twice the following summation by parts formula,

$$\sum_{i=a}^{j} u(i)\Delta v(i) = u(i)v(i)\Big]_a^{j+1} - \sum_{i=a}^{j} v(i+1)\Delta u(i). \tag{22}$$

Refer to (9) and we have, noting that $U_n(0) = U_n(n+1) = 0$,

$$c_j^{(1)} = \frac{1}{\sum\limits_{i=1}^{n} \{y_j(i)\}^2}\Big[-\{\Delta U_n(n+1)\}\sum_{i=1}^{n}\sum_{i=1}^{i} y_j(i) + \\ + \sum_{i=1}^{n}\{\Delta^2 U_n(i)\}\sum_{i=1}^{i}\sum_{i=1}^{i} y_j(i)\Big]$$

$$= \frac{1}{\sum\limits_{i=1}^{n} \{y_j(i)\}^2 (1/n)} \times$$

$$\times\Big[-n\{\Delta U_n(n+1)\}\frac{1}{\sigma_j^2}\sum_{i=1}^{n}\{k(i)n\Delta y_j(i) - k(0)n\Delta y_j(0)\}\frac{1}{n} +$$

$$+\frac{1}{\sigma_j^2}\sum_{i=1}^{n} n^2\Delta^2 U_n(i)\sum_{i=1}^{i}\{k(i)n\Delta y_j(i) - k(0)n\Delta y_j(0)\}\frac{1}{n^2}\Big]. \tag{23}$$

When $p = 1$, our theorem follows from (23). We must apply (22) only once to prove (21) for $c_j^{(2)}$. In the limiting case we have

$$u(x, t) = \sum_{j=1}^{\infty} y_j(x)[a_j \cos \rho_j t + b_j \sin \rho_j t], \tag{24}$$

where $\rho_j$ is a characteristic number and $y_j(x)$ the corresponding characteristic function for the Sturm-Liouville system

$$\frac{d}{dx}\Big[K(x)\frac{dy}{dx}\Big] + \rho^2 y = 0, \quad y(0) = y(1) = 0,$$

and where

$$a_j = \frac{1}{\int\limits_0^1 \{y_j(x)\}^2\,dx}\int\limits_0^1 U(x)y_j(x)\,dx,$$

$$b_j = \frac{1}{\rho_j \int\limits_0^1 \{y_j(x)\}^2\,dx}\int\limits_0^1 V(x)y_j(x)\,dx.$$

We readily prove by integration by parts that $a_j$ and $b_j$ satisfy a relation similar to (21). Moreover, $y_j(x)$ is bounded (6) and consequently $y_j(i)$ which approaches $y_j(x)$ uniformly is bounded.

We are now prepared to state the following theorem which follows from the foregoing considerations, considering (8) as an infinite series with $i = nx$.

THEOREM II. *Series* (8) *converge uniformly in $n$ and $x$ for $0 \leqslant x \leqslant 1$ and for all values of $n$ including the limiting case when $n \to \infty$.*

Now inasmuch as under uniform convergence the limit of the series is the same as the series of the limits, we have the following fundamental theorem, passing directly from (8) and (9).

THEOREM III. *The position of the continuous string is given by* (24). *If $q$ is the abscissa of a point on the string, $q = \mathfrak{M}^{-1}(Mx)$.*

We next make the transformation of independent variable $\mathfrak{M}(q) = Mx$. We have already discussed this transformation. Under it $K(x) = T\delta(q)M^{-2}$. Equation (16) goes into

$$\frac{d^2y}{dq^2} + \frac{\rho^2}{T}\delta(q)y = 0;$$

$u(x, t)$ goes into $w(q, t)$, where $w(q, t)$ is the ordinate of the point on the string with abscissa $q$ at time $t$. Let $U(x)$ go into $X(q)$ and $V(x)$ into $\psi(q)$. We now are in a position to state the following theorem.

THEOREM IV. *Let there be given a stretched string with end-points fixed at $q = 0$ and $q = 1$. Let the tension in the string be uniform and independent of the time. Denote it by $T$. Denote the total mass*

*of the string by M. Denote the density of the string by* $\delta(q)$, *which we assume to be a function with continuous first derivative. Let the string vibrate in a plane with small vibrations. Let* $w(q, t)$ *denote the ordinate of any point of the string, and assume that*

$$w(q, 0) = X(q) \quad \textit{and} \quad w_t(x, 0) = \psi(q).$$

*Assume that* $X(q)$ *is a function with continuous second derivative and* $\psi(q)$ *a function with continuous first derivative at all points of the interval* $0 \leqslant x \leqslant 1$. *Then*

$$w(q, t) = \sum_{j=1}^{\infty} y_j(q)[a_j \cos \rho_j t + b_j \sin \rho_j t],$$

*where* $\rho_j$ *is a characteristic number and* $y_j(x)$ *the corresponding characteristic function for the Sturm-Liouville system*

$$\frac{d^2y}{dq^2} + \frac{1}{T}\delta(q)\rho^2 y = 0, \quad y(0) = y(1) = 0,$$

*and where*

$$a_j = \frac{1}{\int\limits_0^1 \{y_j(q)\}^2 \delta(q)\, dq} \int\limits_0^1 X(q) y_j(q) \delta(q)\, dq,$$

$$b_j = \frac{1}{\rho_j \int\limits_0^1 \{y_j(q)\}^2 \delta(q)\, dq} \int\limits_0^1 \psi(q) y_j(q) \delta(q)\, dq.$$

Proof is immediate from the foregoing considerations and from the fact that under the conditions of our theorem we can construct the functions $K(x)$, $U(x)$, and $V(x)$ and then functions $k(i)$, $U_n(i)$, and $V_n(i)$ which satisfy the requirements that when $i = nx$ and $n \to \infty$, $n^{-2}k(nx) \to K(x)$, $U_n(nx) \to U(x)$, $V_n(nx) \to V(x)$ and

$$n\Delta U_n(nx) \to \frac{dU(x)}{dx}, \quad n\Delta V_n(nx) \to \frac{dV(x)}{dx}, \quad n^2\Delta^2 U_n(nx) \to \frac{d^2U(x)}{dx^2},$$

all uniformly, $0 \leqslant x \leqslant 1$. This is all that is required.

## EXERCISES

1. Study the problem of the weighted vibrating string for more general boundary conditions than those treated in the text.

2. Given a number of hot objects (vases) arranged in line but not touching and radiating heat the one to the other. Set up differential equations for the temperature of the objects and show how the problem can be studied by a linear recurrent relation. Carry through the study in detail for terminal conditions of your own choosing.

## SUGGESTED REFERENCES

BATEMAN, H.: *Partial Differential Equations of Mathematical Physics*, p. 229.

CARSLAW, H. S.: *Theory of Fourier's Series and Integrals*, p. 275.

COURANT, R., and HILBERT, D.: *Methoden der Mathematischen Physik*, **1**, 237.

HOBSON, E. W.: *Proceedings of London Mathematical Society*, series 2, **6**, 378.

LANGER, R. E.: *Amer. Math. Monthly*, vol. **54**, no. 7. (This article was published after the present chapter was in type.)

# XIII

# THE LINEAR RECURRENT RELATION OF THE FIRST ORDER WITH PERIODIC COEFFICIENTS

## 1. Fundamental theorem

WE begin by laying down the following definitions:

*A function $z(i)$ is said to be periodic with the period $\omega$ if*

$$z(i+\omega) \equiv z(i).$$

*It is said to be periodic of the second kind with period $\omega$ and multiplier $\rho$ if*

$$z(i+\omega) \equiv \rho z(i),$$

*where $\rho$ is a constant. It is said to be anti-periodic with period $\omega$ if*

$$z(i+\omega) = -z(i).$$

We shall consider first the homogeneous equation of the first order

$$y(i+1) = p(i)y(i), \tag{1}$$

where $p(i)$ is defined for all integral values of $i$, has the period $\omega$, and does not equal zero at any point.

*Equation* (1) *is called bounded in case every solution remains bounded. It is said to be unbounded in the contrary case.* It is to be remarked that any two solutions are linearly dependent.

Let $y(i)$ be a solution of (1). Then

$$y(i) = y(0)p(0)p(1)...p(i-1), \quad i > 0,$$

$$y(i) = y(0)\frac{1}{p(-1)p(-2)...p(i)}, \quad i < 0.$$

Now due to the periodic character of $p(i)$ we see that

$$p(k)p(k+1)...p(k+\omega-1)$$

is independent of $k$. Denote its value by $L$. We immediately infer the following theorem:

THEOREM I. *Every solution of* (1) *is periodic of the second kind with period $\omega$ and multiplier $L$. The equation is bounded or unbounded according as $|L| \leqslant 1$ or $|L| > 1$. Every solution has the period $\omega$ if $L = 1$. Every solution is anti-periodic with period $\omega$ if $L = -1$.*

For the calculation of $L$ we have the formula

$$L = p(0)p(1)...p(\omega-1). \tag{2}$$

The application of this formula is generally practicable, however, only when $\omega$ is small. If $\omega$ is large, alternative methods must be devised.

## 2. An alternative formula for $L$

Equation (1) of the previous section can be written

$$\Delta y(i)+P(i)y(i) = 0, \tag{3}$$

where $$P(i) = 1-p(i).$$

Then $$L = \{1-P(0)\}\{1-P(1)\}...\{1-P(\omega-1)\}. \tag{4}$$

Let $L_n$ denote the sum of all possible products of

$$P(0),\quad P(1),\quad ...,\quad P(\omega-1),$$

taken $n$ at a time. Then

$$L = L_0-L_1+L_2-L_3+...+(-1)^\omega L_\omega, \quad L_0 = 1. \tag{5}$$

For $L_n$ we can write the formula

$$L_n = \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1}...\sum_{i_n=0}^{i_{n-1}-1} P(i_1)P(i_2)...P(i_n), \quad n > 0.$$

## 3. Important inequality

THEOREM II. *If $P(i)$ is real then*

$$L_j^2 \geqslant \frac{j+1}{j}\,\frac{\omega-j+1}{\omega-j}L_{j-1}\,L_{j+1}, \quad j = 1,2,...,\omega-1. \tag{6}$$

To prove this theorem we consider the equation

$$x^\omega-L_1x^{\omega-1}+L_2x^{\omega-2}-...+(-1)^\omega L_\omega = 0, \quad L_\omega \neq 0.$$

Let $$P(0) = \alpha_1, \quad P(1) = \alpha_2, \quad ..., \quad P(\omega-1) = \alpha_\omega.$$

Now let $j = 1$. Inequality (6) states that

$$(\alpha_1+\alpha_2+...+\alpha_\omega)^2 \geqslant 2\frac{\omega}{\omega-1}(\alpha_1\,\alpha_2+\alpha_1\,\alpha_3+...+\alpha_{\omega-1}\,\alpha_\omega).$$

Assume the contrary and we readily deduce the contradiction

$$(\alpha_1-\alpha_2)^2+(\alpha_1-\alpha_3)^2+...+(\alpha_{\omega-1}-\alpha_\omega)^2 < 0.$$

Now set up the equation whose roots are the reciprocals of the

roots of the equation which we were just discussing and we see that

$$L_{\omega-1}^2 \geqslant 2\frac{\omega}{\omega-1}L_{\omega-2}L_{\omega}. \tag{6'}$$

Now differentiate our equation $(k-2)$ times and we have

$$\omega^{(k-2)}x^{\omega-k+2}-(\omega-1)^{(k-2)}L_1 x^{\omega-k+1}+...+$$
$$+(-1)^{\omega-k}L_{\omega-k}k^{(k-2)}x^2+(-1)^{\omega-k+1}L_{\omega-k+1}(k-1)^{(k-2)}x+$$
$$+(-1)^{\omega-k+2}L_{\omega-k+2}(k-2)^{(k-2)} = 0.$$

Apply (6′) to this, noting that all roots are real. Replace $L_{\omega}$ by $L_{\omega-k+2}(k-2)^{(k-2)}$, $L_{\omega-1}$ by $L_{\omega-k+1}(k-1)^{(k-2)}$, $L_{\omega-2}$ by $L_{\omega-k}k^{(k-2)}$, and $\omega$ by $\omega-k+2$. We arrive at

$$L_{\omega-k+1}^2 \geqslant \frac{k}{k-1}\frac{\omega-k+2}{\omega-k+1}L_{\omega-k}L_{\omega-k+2}.$$

In this we replace $\omega-k+1$ by $j$ and we get (6).

*We remark also the following simple facts: If* $P(i) > 0$ *then* $L_j > 0$. *If* $P(i) \geqslant 0$ *then* $L_j \geqslant 0$, *and in addition if* $L_j = 0$ *then* $L_{j+1} = 0$.

Next we have the following theorem:

THEOREM III. *If* $P(i) \not\equiv 0$ *and* $P(i) \geqslant 0$ *at all points, when*

(a) $1-L_1+L_2-L_3+...+L_{2n} < 1$, *then* $L < 1$,

(b) $1-L_1+L_2-L_3+...-L_{2n-1} > 1$, *then* $L > 1$,

(c) $1-L_1+L_2-L_3+...+L_{2n} < -1$, *then* $L < -1$,

(d) $1-L_1+L_2-L_3+...-L_{2n-1} > -1$, *then* $L > -1$.

To prove this theorem we remark that whenever any one of the relations (a), (b), (c), (d) holds we must have $L_j-L_{j+1} > 0$ at least once in that relation. To show this we treat each case separately, remarking that $L_n \geqslant 0$. In (a), (b), (c) the result is immediate:

(a) $-(L_1-L_2)-(L_3-L_4)-...-(L_{2n-1}-L_{2n}) < 0$,

(b) $-L_1+(L_2-L_3)+...+(L_{2n-2}-L_{2n-1}) > 0$,

(c) $1-(L_1-L_2)-...-(L_{2n-1}-L_{2n}) < -1$,

(d) $(1-L_1)+(L_2-L_3)+...+(L_{2n-2}-L_{2n-1}) > -1$.

We proceed to discuss (d). Assume each parenthesis negative. Then each parenthesis must be numerically less than one. Hence

$L_1 < 2$. But by (6) we have $L_1^2 > 2L_2$. Hence by division $L_1 > L_2$ or $L_1 - L_2 > 0$, and hence $L_2 > L_3$, a contradiction.

Now, by (6), if $L_j - L_{j+1} > 0$ then $L_{j+1} - L_{j+2} \geqslant 0$, and the inequality in question as given in the statement of the theorem is not affected if the series is continued through $L_\omega$ to obtain $L$.

The theorem is useful in determining whether (3) is bounded or unbounded.

We make the following remark: From (6),

$$L_{n+1} < \frac{n}{n+1}\,\frac{\omega-n}{\omega-n+1}\,\frac{L_n^2}{L_{n-1}}.$$

We can thus frequently determine whether $L_{n+1} < L_n$ without actually calculating $L_{n+1}$.

## 4. Upper bounds for $L_n$

The final thought of the last paragraph can be amplified by developing further formulae.

We know that

$$L_n = \sum_{i_1=0}^{\omega-1} \sum_{i_2=0}^{i_1-1} \cdots \sum_{i_n=0}^{i_{n-1}-1} P(i_1)P(i_2)\ldots P(i_n).$$

Suppose $|P(i)| \leqslant M$, then

$$L_n \leqslant M^n \frac{\omega(\omega-1)\ldots(\omega-n+1)}{n!}. \tag{7}$$

Another and frequently more precise bound is obtained as follows.

Let $P \geqslant 0$. Then

$$P(i_1)P(i_2)\ldots P(i_n) \leqslant \frac{1}{n}[\{P(i_1)\}^n + \{P(i_2)\}^n + \ldots + \{P(i_n)\}^n].$$

Hence

$$L_n \leqslant \frac{1}{n} \sum_{i_1=0}^{\omega-1} \sum_{i_2=0}^{i_1-1} \cdots \sum_{i_n=0}^{i_{n-1}-1} [\{P(i_1)\}^n + \{P(i_2)\}^n + \ldots + \{P(i_n)\}^n]$$

$$= \frac{1}{n} \sum [\{P(i_1)\}^n + \{P(i_2)\}^n + \ldots + \{P(i_n)\}^n],$$

where $\sum$ denotes the sum of all expressions

$$[\{P(i_1)\}^n + \{P(i_2)\}^n + \ldots + \{P(i_n)\}^n],$$

where the numbers $i_1, i_2, \ldots, i_n$ are chosen in every possible way from $\omega-1, \omega-2, \ldots, 2, 1, 0$, subject only to the restrictions

$i_1 > i_2 > \ldots > i_n$. The number of ways in which such expressions can be formed is

$$\frac{\omega(\omega-1)\ldots(\omega-n+1)}{n!}.$$

The number of these expressions that will contain $\{P(a)\}^n$ is

$$\frac{(\omega-1)(\omega-2)\ldots(\omega-n+1)}{(n-1)!},$$

where $a = 0, 1, 2, \ldots, \omega-1$. The total coefficient of $\{P(a)\}^n$ is then

$$\frac{(\omega-1)(\omega-2)\ldots(\omega-n+1)}{(n-1)!}.$$

Consequently,

$$L_n \leqslant \frac{(\omega-1)(\omega-2)\ldots(\omega-n+1)}{(n-1)!}\sum_{i=0}^{\omega-1}\{P(i)\}^n. \tag{8}$$

Formula (8) has been derived under the hypothesis that $P \geqslant 0$. However, it still holds if we simply require that $P$ be real, if we also require that $n$ be even.

The bounds that we have derived can be used in the following way. It is desired to tell whether (3) is bounded or unbounded. Suppose that we have calculated $L_1, L_2, \ldots, L_n$ and know that $L < 1$ but do not know whether $L \geqslant -1$ or $L < -1$; that is, we know that

$$L_n > L_{n-1} - L_{n-2} + \ldots \pm L_1 \mp 2,$$

and wish to see if

$$L_{n+1} \leqslant L_n - L_{n-1} + L_{n-2} - \ldots \mp L_1 \pm 2.$$

We may be able to do this by means of (7) or (8) without the necessity of calculating $L_{n+1}$.

## 5. The calculation of $L_n$

We know the formulae

$$L_1 = \sum_{i=0}^{\omega-1} P(i), \tag{9}$$

$$L_2 = \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1} P(i_1)P(i_2), \tag{10}$$

$$L_3 = \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1}\sum_{i_3=0}^{i_2-1} P(i_1)P(i_2)P(i_3), \tag{11}$$

. . . . . . . . . .

Summation by parts may greatly simplify the calculation of $L_2$, $L_3$,.... We find

$$L_2 = \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1} P(i_1)P(i_2) = \Big[\sum_{i=0}^{\omega-1} P(i)\Big]^2 - \sum_{i_1=0}^{\omega-1} P(i_1)\sum_{i_2=0}^{i_1} P(i_2)$$

$$= \Big\{\sum_{i=0}^{\omega-1} P(i)\Big\}^2 - \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1} P(i_1)P(i_2) - \sum_{i=0}^{\omega-1}\{P(i)\}^2.$$

Transposing and dividing by 2 we have

$$L_2 = \tfrac{1}{2}\Big[\sum_{i=0}^{\omega-1} P(i)\Big]^2 - \tfrac{1}{2}\sum_{i=0}^{\omega-1}\{P(i)\}^2. \tag{12}$$

A similar process yields

$$L_3 = \tfrac{1}{3}\Big[\sum_{i=0}^{\omega-1} P(i)\Big]\Big[\tfrac{1}{2}\Big\{\sum_{i=0}^{\omega-1} P(i)\Big\}^2 - \tfrac{1}{2}\sum_{i=0}^{\omega-1}\{P(i)\}^2\Big] -$$

$$-\tfrac{1}{3}\sum_{i_1=0}^{\omega-1} P(i_1)\sum_{i_2=0}^{i_1-1}\{P(i_2)\}^2 - \tfrac{1}{3}\sum_{i_1=0}^{\omega-1}\{P(i_1)\}^2\sum_{i_2=0}^{i_1-1} P(i_2). \tag{13}$$

From (9), (12), and (13),

$$L_2 = \tfrac{1}{2}L_1^2 - \tfrac{1}{2}\sum_{i=0}^{\omega-1}\{P(i)\}^2, \tag{14}$$

$$L_3 = \tfrac{1}{3}L_1 L_2 - \tfrac{1}{3}\sum_{i_1=0}^{\omega-1} P(i_1)\sum_{i_2=0}^{i_1-1}\{P(i_2)\}^2 - \tfrac{1}{3}\sum_{i_1=0}^{\omega-1}\{P(i_1)\}^2\sum_{i_2=0}^{i_1-1} P(i_2). \tag{15}$$

## 6. An example

We wish to determine whether the following equation is bounded or unbounded:

$$\Delta y(i) + P(i)y(i) = 0, \tag{16}$$

$P(i) = \lambda(10)^{-i}$, $0 \leqslant i \leqslant \omega-1$. Also $P(i)$ is periodic with period $\omega$. Here $\lambda > 0$ is an arbitrary parameter to be discussed later.

We shall consider (16) for large values of $\omega$.

It is immediate from (4) that $0 < L < 1$ if $\lambda < 1$. Equation (16) is bounded under this circumstance. It also is immediate that $p(i) = 1 - P(i) = 0$ if $\lambda(10)^{-i} = 1$. We consequently shall assume that $\lambda \neq 10^i$ for any value of $i$. We shall proceed to the calculation of $L_1$, $L_2$, and $L_3$ for other values of $\lambda$.

$$\sum_{i_1=0}^{i-1} P(i_1) = \lambda\sum_{i_1=0}^{i-1}(0{\cdot}1)^{i_1} = \lambda\tfrac{10}{9}[1-(0{\cdot}1)^i], \tag{17}$$

$$\sum_{i_1=0}^{i-1}\{P(i_1)\}^2 = \lambda^2\sum_{i_1=0}^{i-1}(0{\cdot}01)^{i_1} = \lambda^2\tfrac{100}{99}[1-(0{\cdot}01)^i], \tag{18}$$

$$\sum_{i_1=0}^{i-1} P(i_1) \sum_{i_2=0}^{i_1-1} \{P(i_2)\}^2 = \lambda^3 \sum_{i_1=0}^{i-1} \tfrac{100}{99}[(0{\cdot}1)^{i_1} - (0{\cdot}001)^{i_1}]$$
$$= \lambda^3 \tfrac{100}{99}[\tfrac{10}{9}\{1-(0{\cdot}1)^i\} - \tfrac{1000}{999}\{1-(0{\cdot}001)^i\}], \quad (19)$$

$$\sum_{i_1=0}^{i-1} \{P(i_1)\}^2 \sum_{i_2=0}^{i_1-1} P(i_2) = \lambda^3 \sum_{i_1=0}^{i-1} [\tfrac{10}{9}\{(0{\cdot}01)^{i_1} - (0{\cdot}001)^{i_1}\}]$$
$$= \lambda^3 \tfrac{10}{9}[\tfrac{100}{99}\{1-(0{\cdot}01)^i\} - \tfrac{1000}{999}\{1-(0{\cdot}001)^i\}]. \quad (20)$$

Since $\omega$ is large we have from these formulae and from (9), (14), and (15) the following approximate results:

$$L_1 \simeq \lambda \tfrac{10}{9},$$
$$L_2 \simeq \tfrac{1}{2}\lambda^2[\tfrac{100}{81} - \tfrac{100}{99}],$$
$$L_3 \simeq \tfrac{1}{6}\cdot\tfrac{10}{9}[\tfrac{100}{81} - \tfrac{100}{99}] - \tfrac{1}{3}[\tfrac{10}{9}(\tfrac{100}{99} - \tfrac{1000}{999}) + \tfrac{100}{99}(\tfrac{10}{9} - \tfrac{1000}{999})].$$

These results can be rewritten:

$$L_1 \simeq \lambda(1{\cdot}111111),$$
$$L_2 \simeq \lambda^2(0{\cdot}112233),$$
$$L_3 \simeq \lambda^3(0{\cdot}001123).$$

We now refer to Theorem III ($a$) and ($d$) and see that, if the following two inequalities are simultaneously satisfied, $|L| < 1$, and hence that (16) is bounded:

$$\lambda(1{\cdot}111111) - \lambda^2(0{\cdot}112233) > 0,$$
$$2 - \lambda(1{\cdot}111111) + \lambda^2(0{\cdot}112233) - \lambda^3(0{\cdot}001123) > 0.$$

The first of these inequalities is satisfied if $\lambda < 9{\cdot}9$. The second inequality is satisfied for various values of $\lambda$, in particular if

$$9 \leqslant \lambda \leqslant 10.$$

From Theorem III ($b$) equation (16) is unbounded if

$$\lambda^2(0{\cdot}001123) - \lambda(0{\cdot}112233) + 1{\cdot}111111 < 0,$$

which is true if $\lambda$ lies between the real roots of the corresponding quadratic equation.

It will be observed that the results obtained in the above discussion have to do with the value of an infinite product. The method used may be of interest in treating certain infinite products which arise elsewhere than in the theory of finite differences.

## 7. The non-homogeneous equation

Consider the equation

$$y(i+1)-p(i)y(i) = r(i), \tag{21}$$

where $p(i)$ and $r(i)$ have the period $\omega$ and $p(i) \neq 0$ at any point.

Denote by $y_1(i)$ a particular non-vanishing solution of the reduced equation

$$y(i+1)-p(i)y(i) = 0. \tag{22}$$

We know that $y_1(i)$ is periodic of the second kind with period $\omega$ and multiplier $L$.

The general solution of (1) is given by

$$y(i) = y_1(i)\left[\sum_{i=0}^{i-1} \frac{r(i)}{y_1(i+1)}+C\right], \tag{23}$$

where $C$ is arbitrary.

Now form the expression

$$y(i+\omega)-y(i) \equiv Ly_1(i)\sum_{i_1=0}^{i+\omega-1} \frac{r(i)}{y_1(i+1)}+LCy_1(i)- -y_1(i)\sum_{i=0}^{i-1} \frac{r(i)}{y_1(i+1)}-Cy_1(i).$$

But $$\sum_{i=0}^{i+\omega-1} \frac{r(i)}{y_1(i+1)} = \sum_{i=0}^{\omega-1} \frac{r(i)}{y_1(i+1)}+\frac{1}{L}\sum_{i=0}^{i-1} \frac{r(i)}{y_1(i+1)}.$$

Hence

$$y(i+\omega)-y(i) = Ly_1(i)\sum_{i=0}^{\omega-1} \frac{r(i)}{y_1(i+1)}+(L-1)Cy_1(i).$$

For brevity we write

$$y(i+\omega)-y(i) = Ky_1(i). \tag{24}$$

Here $K$ is constant.

If $L \neq 1$, then $K$ can be given any arbitrary value by a proper choice of $C$ and conversely. Since to every $C$ there corresponds a unique solution of equation (21) we conclude that, if $L \neq 1$, to every $K$ there corresponds a unique solution of (21). Sum both sides of (24) for $i = i,\ i+\omega,\ i+2\omega,\ldots,\ i+(n-1)\omega$. We get

$$y(i+n\omega) = K\sum_{j=0}^{n-1} y_1(i+j\omega)+y(i).$$

Now $y_1(i+\omega) = Ly_1(i)$. Hence

$$y(i+n\omega) = Ky_1(i)\sum_{j=0}^{n-1} L^j+y(i).$$

We draw the following conclusions which are formulated as a theorem.

THEOREM IV. *Let $y$ be that solution of* (21) *given by* (23), *and let*

$$K = (L-1)C+L\sum_{i=0}^{\omega-1}\frac{r(i)}{y_1(i+1)}.$$

*If $K \neq 0$ and $L < 1$ then $y$ is bounded.*

*If $K \neq 0$ and $L \geqslant 1$ then $y$ is unbounded.*

*If $L \neq 1$ there is always one and only one periodic solution, namely that one for which $K = 0$.*

*If $L = 1$ there is in general no periodic solution. However, all solutions will be periodic if $K$, which is now independent of $C$, is zero.*

## 8. Right-hand member periodic of second kind

Consider
$$\Delta y(i) = r(i), \quad i \geqslant 0, \tag{25}$$
where $r(i)$ is periodic of the second kind, with period $\omega$ and multiplier $\rho \neq 1$.

The general solution of (25) is

$$y(i) = \sum_{i=0}^{i-1} r(i)+C, \quad C = y(0), \; i > 0. \tag{26}$$

Assume the existence of a solution $y$ periodic of the second kind with period $\omega$ and multiplier $\rho$. Then

$$y(i+\omega) = \rho y(i) = \rho\sum_{i=0}^{i-1} r(i)+\rho C, \quad i > 0,$$

and
$$y(i+\omega) = \sum_{i=0}^{i+\omega-1} r(i)+C = \sum_{i=0}^{\omega-1} r(i)+\rho\sum_{i=0}^{i-1} r(i)+C.$$

Equating these values we have

$$C[\rho-1] = \sum_{i=0}^{\omega-1} r(i). \tag{27}$$

This is necessary and sufficient for the existence of $y$ as assumed. There is clearly a unique determination of $C$ satisfying (27) and hence one and only one solution of (25) periodic of the second

kind with period $\omega$ and multiplier $\rho$. The form of the general solution is $y(i) = f(i)+C$, where $C$ is an arbitrary constant and $f(i)$ periodic of the second class.

### EXERCISES

1. Discuss $y(i+1) = p(i)y(i)$, where $p(i)$ is periodic of the second kind with multiplier $\rho \neq 1$.

2. Discuss

(a) $\Delta y(i)+\lambda \sin^2\frac{\pi i}{\omega}y(i) = 0;$

(b) $\Delta y(i)+9i10^{-i}y(i) = 0;$

(c) $\Delta y(i)+\lambda\left(2-\sin\frac{\pi i}{\omega}\right)y(i) = 0.$

XIV

# THE LINEAR RECURRENT RELATION OF THE SECOND ORDER WITH PERIODIC COEFFICIENTS

## 1. Periodic solutions

CONSIDER the equation

$$y(i+2)+p(i)y(i+1)+q(i)y(i) = 0, \tag{1}$$

where $p(i)$ and $q(i)$ are real and defined for all values of $i$ each with the period $\omega$, and $q(i) \neq 0$ at any point. We ask if solutions exist satisfying identically the relation

$$y(i+\omega) = y(i), \tag{2}$$

and also if solutions exist satisfying identically

$$y(i+\omega) = -y(i). \tag{3}$$

A solution satisfying (2) is called periodic and a solution satisfying (3) anti-periodic.

The following theorem is immediate.

THEOREM 1. *A necessary and sufficient condition that a solution $y(i)$ of (1) be periodic is that*

$$\begin{aligned} y(a) &= y(a+\omega), \\ \Delta y(a) &= \Delta y(a+\omega), \end{aligned} \tag{4}$$

*where $a$ is any particular point, and that a solution $y(i)$ be anti-periodic is that*

$$\begin{aligned} y(a) &= -y(a+\omega), \\ \Delta y(a) &= -\Delta y(a+\omega). \end{aligned} \tag{5}$$

## 2. The characteristic equation

We propose the question: *Does a solution $y(i)$ of (1) exist such that*

$$y(i+\omega) = \rho y(i), \tag{6}$$

*where $\rho$ is a constant?*

Let $y_1(i)$ and $y_2(i)$ be two linearly independent solutions. There exist two linearly independent solutions $\bar{y}_1(i)$ and $\bar{y}_2(i)$ such that

$$\begin{aligned} \bar{y}_1(i+\omega) &\equiv y_1(i), \\ \bar{y}_2(i+\omega) &\equiv y_2(i). \end{aligned} \tag{7}$$

This is true, since if $a$ is any point the equations

$$\bar{y}_1(a+\omega) = y_1(a), \qquad \bar{y}_2(a+\omega) = y_2(a),$$
$$\bar{y}_1(a+\omega+1) = y_1(a+1), \qquad \bar{y}_2(a+\omega+1) = y_2(a+1),$$

serve to determine $\bar{y}_1(i)$ and $\bar{y}_2(i)$ uniquely.

Now any two solutions such as $y_1(i)$ and $y_2(i)$ are linear combinations of $\bar{y}_1(i)$ and $\bar{y}_2(i)$. Consequently, using (7),

$$\begin{aligned} y_1(i+\omega) &= \alpha_{11}\, y_1(i)+\alpha_{12}\, y_2(i), \\ y_2(i+\omega) &= \alpha_{21}\, y_1(i)+\alpha_{22}\, y_2(i). \end{aligned} \tag{8}$$

Assume now the existence of a solution $y_3(i) \not\equiv 0$ such that

$$y_3(i+\omega) = \rho y_3(i). \tag{9}$$

Then

$$\begin{aligned} y_3(i+\omega) &= \beta_1\, y_1(i+\omega)+\beta_2\, y_2(i+\omega) \\ &= \beta_1\{\alpha_{11}\, y_1(i)+\alpha_{12}\, y_2(i)\}+\beta_2\{\alpha_{21}\, y_1(i)+\alpha_{22}\, y_2(i)\} \\ &= \rho y_3(i) = \rho\{\beta_1\, y_1(i)+\beta_2\, y_2(i)\}. \end{aligned}$$

Transposing, we have

$$\{\beta_1(\alpha_{11}-\rho)+\beta_2\,\alpha_{21}\}y_1(i)+\{\beta_1\,\alpha_{12}+\beta_2(\alpha_{22}-\rho)\}y_2(i) = 0.$$

Since $y_1(i)$ and $y_2(i)$ are linearly independent,

$$\beta_1(\alpha_{11}-\rho)+\beta_2\,\alpha_{21} = 0,$$
$$\beta_1\,\alpha_{12}+\beta_2(\alpha_{22}-\rho) = 0.$$

But $\beta_1$ and $\beta_2$ are not both zero since $y_3(i) \not\equiv 0$. Hence

$$D_1 \equiv (\alpha_{11}-\rho)(\alpha_{22}-\rho)-\alpha_{12}\,\alpha_{21} = 0. \tag{10}$$

This has been derived as a necessary condition. However, if $\rho$ is so chosen that $D_1 = 0$ we can retrace steps and see that a solution not identically zero satisfying the relation (6) actually does exist.

From (8)

$$\begin{vmatrix} y_1(i+\omega+1) & y_1(i+\omega) \\ y_2(i+\omega+1) & y_2(i+\omega) \end{vmatrix} = \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} \cdot \begin{vmatrix} y_1(i+1) & y_1(i) \\ y_2(i+1) & y_2(i) \end{vmatrix}.$$

But by use of (1)

$$\begin{vmatrix} y_1(i+2) & y_1(i+1) \\ y_2(i+2) & y_2(i+1) \end{vmatrix} = q(i) \begin{vmatrix} y_1(i+1) & y_1(i) \\ y_2(i+1) & y_2(i) \end{vmatrix}.$$

Hence

$$\begin{vmatrix} y_1(i+\omega+1) & y_1(i+\omega) \\ y_2(i+\omega+1) & y_2(i+\omega) \end{vmatrix} = q(i)q(i+1)...q(i+\omega-1)\begin{vmatrix} y_1(i+1) & y_1(i) \\ y_2(i+1) & y_2(i) \end{vmatrix}.$$

Hence $$\begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} = q(i)q(i+1)...q(i+\omega-1).$$

We denote this by $M$ and rewrite (10) as follows:

$$\begin{aligned} \rho^2-(\alpha_{11}+\alpha_{22})\rho+M &= 0, \\ M &= q(i)q(i+1)...q(i+\omega-1). \end{aligned} \tag{11}$$

Regard (11) as an equation in $\rho$. We call it the *characteristic equation*. There will then always exist at least one solution of (1), not identically zero, satisfying (6). If equation (11) has two distinct roots $\rho_1$ and $\rho_2$ there exist at least two† solutions $y_4(i)$ and $y_5(i)$, not either identically zero, which satisfy respectively the relations

$$y_4(i+\omega) = \rho_1 y_4(i) \quad \text{and} \quad y_5(i+\omega) = \rho_2 y_5(i).$$

Moreover $y_4(i)$ and $y_5(i)$ are linearly independent for, assuming the contrary, there exist two constants $\mu_1$ and $\mu_2$ not both zero such that

$$\mu_1 y_4(i)+\mu_2 y_5(i) \equiv 0. \tag{12}$$

From (12), $$\mu_1 y_4(i+\omega)+\mu_2 y_5(i+\omega) \equiv 0.$$

Hence $$\mu_1 \rho_1 y_4(i)+\mu_2 \rho_2 y_5(i) \equiv 0. \tag{13}$$

For definiteness assume $\mu_1 \neq 0$. Eliminate $y_5(i)$ from (12) and (13) and we get

$$\mu_1(\rho_1-\rho_2)y_4(i) \equiv 0.$$

But $\mu_1 \neq 0$, $y_4(i) \not\equiv 0$. Hence $\rho_1 = \rho_2$. This is a contradiction. Hence $y_4(i)$ and $y_5(i)$ are linearly independent.

We shall now prove the following theorem.

THEOREM II. *The characteristic equation is independent of the particular fundamental system of solutions chosen.*

Consider a second fundamental system of solutions $y_7(i)$ and $y_8(i)$. Then

$$\begin{aligned} y_7(i+\omega) &= B_{11} y_7(i)+B_{12} y_8(i), \\ y_8(i+\omega) &= B_{21} y_7(i)+B_{22} y_8(i). \end{aligned}$$

† These solutions may be imaginary.

The characteristic equation is

$$D_2 = (B_{11}-\rho)(B_{22}-\rho)-B_{12}B_{21} = 0.$$

But

$$y_7(i) = L_{11}y_1(i)+L_{12}y_2(i),$$
$$y_8(i) = L_{21}y_1(i)+L_{22}y_2(i),$$

where $L_{11}L_{22}-L_{12}L_{21} \neq 0$. Hence

$$y_7(i+\omega) = B_{11}[L_{11}y_1(i)+L_{12}y_2(i)]+B_{12}[L_{21}y_1(i)+L_{22}y_2(i)].$$

On the other hand one can write

$$\begin{aligned} y_7(i+\omega) &= L_{11}y_1(i+\omega)+L_{12}y_2(i+\omega) \\ &= L_{11}[\alpha_{11}y_1(i)+\alpha_{12}y_2(i)]+L_{12}[\alpha_{21}y_1(i)+\alpha_{22}y_2(i)]. \end{aligned}$$

Equate these two right-hand members and collect:

$$\begin{aligned} &[(B_{11}L_{11}+B_{12}L_{21})-(L_{11}\alpha_{11}+L_{12}\alpha_{21})]y_1(i)+ \\ &\qquad +[(B_{11}L_{12}+B_{12}L_{22})-(L_{11}\alpha_{12}+L_{12}\alpha_{22})]y_2(i) = 0. \end{aligned}$$

Here the coefficients must be zero as $y_1(i)$ and $y_2(i)$ are linearly independent; that is,

$$B_{11}L_{11}+B_{12}L_{21} = L_{11}\alpha_{11}+L_{12}\alpha_{21},$$
$$B_{11}L_{12}+B_{12}L_{22} = L_{11}\alpha_{12}+L_{12}\alpha_{22}.$$

Similarly

$$B_{21}L_{11}+B_{22}L_{21} = L_{21}\alpha_{11}+L_{22}\alpha_{21},$$
$$B_{21}L_{12}+B_{22}L_{22} = L_{21}\alpha_{12}+L_{22}\alpha_{22}.$$

If these four equations are solved for $B_{11}$, $B_{12}$, $B_{21}$, $B_{22}$ and substitution made in $D_2$, we find $D_2 = D_1$, which we desired to prove.

## 3. Simple theorems

The following four theorems are so easy that proofs are omitted.

THEOREM III. *If $\rho_1$ and $\rho_2$ are real then $y_4(i)$ and $y_5(i)$ are real but for possible constant multipliers.*

THEOREM IV. *When $\rho_1$ and $\rho_2$ are imaginary $y_4(i)$ and $y_5(i)$ are conjugate but for a possible constant multiplier.*

THEOREM V. *When $\rho_1$ and $\rho_2$ are imaginary and $y_4(i)$ and $y_5(i)$ taken conjugate, in order for a solution $c_1y_4(i)+c_2y_5(i)$ to be real it is necessary and sufficient that $c_1$ and $c_2$ be conjugate.*

THEOREM VI. *If* $\rho_1$ *and* $\rho_2$ *are equal they are real, and a solution satisfying the relation,* $y(i+\omega) \equiv \rho y(i)$, *is real but for a constant multiplier.*

## 4. The general solution

Suppose that $\rho_1 = \rho_2$. Denote the common value by $\rho_1$. Let $y_6(i)$ be a solution such that

$$y_6(i+\omega) = \rho_1 y_6(i).$$

Let $y_7(i)$ be a particular solution linearly independent of $y_6(i)$, then

$$y_6(i+\omega) = \rho_1 y_6(i),$$

$$y_7(i+\omega) = c_{21} y_6(i)+c_{22} y_7(i).$$

The characteristic equation is

$$(\rho_1-\rho)(c_{22}-\rho) = 0.$$

This must have $\rho_1$ a double root. Hence $c_{22} = \rho_1$. In other words

$$y_7(i+\omega) = c_{21} y_6(i)+\rho_1 y_7(i).$$

By division

$$\frac{y_7(i+\omega)}{y_6(i+\omega)} = \frac{y_7(i)}{y_6(i)}+\frac{c_{21}}{\rho_1} \quad \text{when} \quad y_6(i) \neq 0.$$

It results that $\dfrac{y_7(i)}{y_6(i)}-\dfrac{c_{21}}{\rho_1 \omega} i$ is single-valued and has the period $\omega$. Call it $\phi(i)$. If $y_6(\bar{i}) = 0$ then $\phi(i)$ is not defined when

$$i = \bar{i}+n\omega, \quad n = \pm 1, \pm 2,\ldots.$$

We have
$$y_7(i) = \phi(i)y_6(i)-i\frac{c_{21}}{\rho_1 \omega} y_6(i)$$

when $y_6(i) \neq 0$. We accordingly write

$$y_7(i) = ciy_6(i)+y_8(i),$$

where $c = c_{21}/\rho_1 \omega$ and $y_8(i) = \phi(i)y_6(i)$ when $y_6(i) \neq 0$. When $y_6(i) = 0$ then $y_8(i) = y_7(i)$. We know that $\phi(i)y_6(i)$ is periodic of the second kind with multiplier $\rho_1$ when $y_6(i) \neq 0$. But $y_7(i+\omega) = \rho y_7(i)$ at the points where $y_6(i) = 0$. Hence $y_8(i)$ is periodic of the second kind with multiplier $\rho_1$ for all values of $i$.

Now $y_6(i)$ and $y_7(i)$ constitute a fundamental system. Consequently the general solution when $\rho_1 = \rho_2$ is of the form

$$y = \bar{c}_1 y_6(i) + c_2\{ciy_6(i) + y_8(i)\}.$$

We write this $$y = c_1 f_1(i) + c_2 i f_2(i),$$

where $f_1(i)$ and $f_2(i)$ are each periodic of the second kind with period $\omega$ and multiplier $\rho_1$.

In case $\rho_1 \neq \rho_2$ the general solution is given by

$$y = c_1 y_1(i) + c_2 y_2(i),$$

where $y_1(i)$ and $y_2(i)$ are periodic of the second kind with multipliers $\rho_1$ and $\rho_2$ respectively.

We remark at this point that any function $F(i)$ periodic of the second kind with multiplier $\rho$ and period $\omega$ can be written

$$F(i) = \rho^{i/\omega} f(i),$$

where $f(i)$ is periodic.

We consequently can write the general solution as follows:

$$y = c_1 \rho_1^{i/\omega}\phi(i) + c_2 i \rho_1^{i/\omega}\psi(i), \quad \rho_1 = \rho_2,$$

$$y = c_1 \rho_1^{i/\omega}\chi(i) + c_2 \rho_2^{i/\omega}\Omega(i), \quad \rho_1 \neq \rho_2,$$

where $\phi(i)$, $\psi(i)$, $\chi(i)$, and $\Omega(i)$ are periodic with period $\omega$.

## 5. Sturm's normal form and the functions $\lambda_j(x)$

We now further particularize equation (1). We consider

$$\Delta[K(i,\lambda)\Delta y(i)] - G(i,\lambda)y(i+1) = 0, \tag{14}$$

where both $K$ and $G$ have the period $\omega$ as functions of $\lambda$ and are continuous. We assume moreover that $K(i,\lambda) > 0$ is bounded and never decreases as $\lambda$ increases, also that $G(i,\lambda)$ increases from $-\infty$ to $\infty$ as $\lambda$ increases.

Now $$q(i)q(i+1)...q(i+\omega-1) = \frac{K(i,\lambda)}{K(i+\omega,\lambda)} = 1,$$

and (11) reduces to

$$\rho^2 - (\alpha_{11} + \alpha_{22})\rho + 1 = 0. \tag{15}$$

Let us refer to Theorems III and IV of Chapter X. We now define the function $y(i)$ as a function $y(x)$ of a continuous argument, namely as the function defined by the broken-line graph

of $y(i)$. We shall denote by $y'_+(x)$ the forward derivative of $y(x)$ and by $y'_-(x)$ the backward derivative of $y(x)$. Theorems III and IV state that there exist values $\lambda_j$ such that when $\lambda = \lambda_j$ all solutions of (14), not identically zero, vanishing at $x$ vanish at $x+\omega$ also with exactly $j$ zeros in between. If $x$ is an integer $j = 0, 1,..., \omega-2$; if $x$ is not an integer $j = 0, 1,..., \omega-1$.

A value $\lambda_j$ is a function of the beginning point $x$. We write $\lambda_j(x)$.

THEOREM VII. *$\lambda_j(x)$ is a continuous function for all values of $x$ when $j = 0, 1,..., \omega-2$. Moreover $\lambda_{\omega-1}(x)$ is a continuous function except where $x$ is integral and $\lambda_{\omega-1}(x)$ becomes negatively infinite as $x$ approaches any integer.*

Let $\xi$ be any particular value of $x$ and let $y_\xi(x)$ be a particular solution not identically zero such that $y_\xi(\xi) = 0$. Let $c$ be that integer such that $c < \xi \leqslant c+1$.

Let $\zeta$ be a second real number and let $|\zeta-\xi| < \delta$. Let $y_\zeta(x)$ be a particular solution not identically zero such that $y_\zeta(\zeta) = 0$. Suppose $\zeta < c+1$. Then, let $y'_{\zeta-}(\zeta) = y'_{\xi-}(\xi)$. If, on the other hand, $\zeta > c+1$ let $y'_{\zeta+}(\zeta) = y'_{\xi+}(\xi)$. In the first instance $|y_\xi(c)-y_\zeta(c)| < \epsilon$ and $|y_\xi(c+1)-y_\zeta(c+1)| < \epsilon$, where $\epsilon > 0$, if $\delta$ is sufficiently small; and in the second instance

$$|y_\xi(c+1)-y_\zeta(c+1)| < \epsilon \quad \text{and} \quad |y_\xi(c+2)-y_\zeta(c+2)| < \epsilon.$$

But by solving (14) for successive values we obtain $y_\xi(c+\omega)$ as a polynomial in $y_\xi(c)$ and $y_\xi(c+1)$ with coefficients continuous functions of $\lambda$. Moreover, $y_\zeta(c+\omega)$ is the same polynomial function of $y_\zeta(c)$ and $y_\zeta(c+1)$. Similarly,

$$y_\xi(c+\omega+1),\, y_\xi(c+\omega+2) \quad \text{and} \quad y_\zeta(c+\omega+1),\, y_\zeta(c+\omega+2)$$

are given by the same polynomial in

$$y_\xi(c+1),\, y_\xi(c+2) \quad \text{and} \quad y_\zeta(c+1),\, y_\zeta(c+2),$$

respectively. Consequently, if $\epsilon$ is sufficiently small,

$$|y_\xi(c+\omega)-y_\zeta(c+\omega)| < \eta,$$
$$|y_\xi(c+\omega+1)-y_\zeta(c+\omega+1)| < \eta,$$
$$|y_\xi(c+\omega+2)-y_\zeta(c+\omega+2)| < \eta,$$

where $\eta$ is arbitrarily small. Now $y_\xi(x)$ has a zero at $\xi+\omega$ and crosses the axis at that point. Consequently, $y_\zeta(x)$ has a zero arbitrarily close to $\xi+\omega$ crossing the axis. Now as $\lambda$ increases zeros of $y_\zeta(x)$ move continuously to the right and as $\lambda$ decreases continuously to the left, actually reaching the point $\zeta+\omega$ if $j < \omega-1$ or if $\zeta$ is non-integral. Consequently, in these cases, given a $\sigma > 0$ it is possible to find a $\delta$ such that when $|\xi-\zeta| < \delta$

$$|\lambda_j(\xi)-\lambda_j(\zeta)| < \sigma.$$

It is well to remark at this point that the roots of $y_\zeta(x)$ are in no way changed by the assignment of a forward, or if you please, backward derivative at $\zeta$. The solution is determined, but for a constant multiplier, by the fact that $y_\zeta(\zeta) = 0$.

The exceptional situation in the case of $\lambda_{\omega-1}(x)$ arises from the fact that if $\xi$ is an integer no value of $\lambda_{\omega-1}(\xi)$ exists but that as $\lambda \to -\infty$ the root on the interval $\xi+\omega < x < \xi+\omega+1$ approaches arbitrarily close to $\xi+\omega$.

## 6. The maxima and minima of the functions $\lambda_j(x)$

We shall prove the following theorem:

THEOREM VIII. *The maxima and minima of the functions $\lambda_j(x)$ are values for which a solution exists satisfying identically either* (2) *or* (3).

Under maximum (minimum) we include the case that the function is a constant or has a constant value over a neighbourhood.

Suppose $\lambda_j(a)$ a maximum. We shall prove that $\rho = \pm 1$. Assume that this is not the case, then $\rho_1 \neq \rho_2$ since $\rho_1 \rho_2 = 1$.

As previously, denote by $y_4(x)$ and $y_5(x)$ solutions such that

$$\begin{aligned} y_4(x+\omega) &\equiv \rho_1 y_4(x), \\ y_5(x+\omega) &\equiv \rho_2 y_5(x). \end{aligned} \tag{16}$$

Then

$$y_4(x)y_5(x+\omega)-y_5(x)y_4(x+\omega) = (\rho_2-\rho_1)y_4(x)y_5(x). \tag{17}$$

Let $y_a(x)$ be a solution, not identically zero, such that

$$y_a(a) = y_a(a+\omega) = 0$$

when $\lambda = \lambda_j(a)$. Then

$$y_a(x) = C_1 y_4(x)+C_2 y_5(x).$$

Hence

$$0 = C_1 y_4(a)+C_2 y_5(a),$$

$$0 = \rho_1 C_1 y_4(a)+C_2 \rho_2 y_5(a).$$

As $C_1$ and $C_2$ are not both zero,

$$(\rho_1-\rho_2)y_4(a)y_5(a) = 0.$$

But

$$\rho_1 \neq \rho_2.$$

Hence

$$y_4(a)y_5(a) = 0.$$

But $y_4(a)$ and $y_5(a)$ are not both zero, since $y_4(x)$ and $y_5(x)$ are linearly independent, and in some neighbourhood of $a$, excluding $a$ itself, neither is zero. Suppose for definiteness $y_4(a) = 0$. Now $y_4(x)$ is essentially real, that is, real but for a constant multiplier. Write $y_4(x) = y_p(x)+\sqrt{(-1)}y_q(x)$, where $y_p(x)$ and $y_q(x)$ are real. Then $y_p(x)$ and $y_q(x)$ are also solutions of (1). But $y_p(a) = y_q(a) = 0$. Hence $y_p(x)$ and $y_q(x)$ are proportional as is the case with all solutions which vanish at the same point. Consequently, $y_4(x) \equiv cy_p(x)$. Let us then so choose $c$ as to make $y_4(x)$ real. It changes sign at $a$. Also $y_5(a) \neq 0$.

Now let $y_1(x)$ and $y_2(x)$ be any two real linearly independent solutions. Then

$$y_1(x) = k_{11} y_4(x)+k_{12} y_5(x),$$

$$y_2(x) = k_{21} y_4(x)+k_{22} y_5(x);$$

$$U(x) \equiv y_1(x)y_2(x+\omega)-y_2(x)y_1(x+\omega) \tag{18}$$

$$= (k_{11} k_{22}-k_{21} k_{12})\{y_4(x)y_5(x+\omega)-y_5(x)y_4(x+\omega)\}.$$

It results from (17) that $U(x)$ changes sign at $a$, that is, if $\delta > 0$ is sufficiently small,

$$U(a+\delta)U(a-\delta) < 0.$$

Under the assumption that $\lambda_j(a)$ is a maximum we shall derive a contradiction to this statement.

Let $y_a(x)$ be a solution such $y_a(a) = 0$. Choose points $a-\delta$ and $a+\delta$ and two real solutions, not identically zero, $y_{a-\delta}(x)$ and $y_{a+\delta}(x)$, where $y_{a-\delta}(a-\delta) = 0$ and $y_{a+\delta}(a+\delta) = 0$, with forward

derivatives at $a-\delta$ and $a+\delta$ respectively equal to the forward derivative of $y_a(x)$ at $a$. When $\lambda = \lambda_j(a)$ the situation is somewhat as here illustrated.


We conclude immediately that

$$y_{a-\delta}(a\pm\delta)y_{a+\delta}(a+\omega\pm\delta)-y_{a-\delta}(a+\omega\pm\delta)y_{a+\delta}(a\pm\delta) \geqslant 0. \quad (19)$$

As these are two linearly independent solutions similar to $y_1(x)$ and $y_2(x)$ we see by (18) that

$$U(a+\delta)U(a-\delta) \geqslant 0, \quad (20)$$

which is the desired contradiction.

Formal changes only are necessary to treat the minima of $\lambda_j(x)$.

THEOREM IX. *If when* $\lambda = \Lambda$, $\rho_1 = \rho_2$, *and a solution not identically zero satisfying* (2) *or* (3) *has a zero at* $a$, *then* $\Lambda$ *is necessarily a maximum or minimum of* $\lambda_j(x)$ *and* $\Lambda = \lambda_j(a)$.

The terms maximum and minimum are used as heretofore.

When $\rho_1 = \rho_2$ there exists a real solution, not identically zero, which we shall call $y_7(x)$ such that $y_7(x+\omega) = \rho y_7(x)$ and a linearly independent real solution $y_8(x)$ where, as we have seen,

$$y_8(x+\omega) = C_{21}y_7(x)+\rho y_8(x).$$

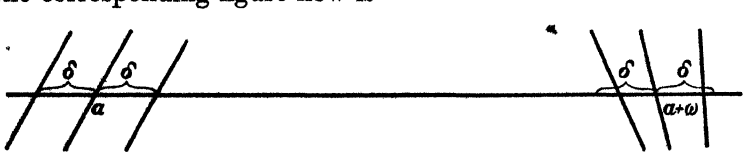
Then

$$U(x) = y_7(x)y_8(x+\omega)-y_7(x+\omega)y_8(x) = C_{21}[y_7(x)]^2.$$

This never changes sign and only vanishes when $y_7(x) = 0$. Consequently,

$$U(a+\delta)U(a-\delta) > 0. \quad (21)$$

Assume $\Lambda$ not an extreme of $\lambda_j(x)$. Refer to the determinant (19). The corresponding figure now is



From this we conclude that $U$, which is a real constant times

$$y_{a-\delta}(x)y_{a+\delta}(x+\omega)-y_{a-\delta}(x+\omega)y_{a+\delta}(x) \tag{22}$$

does not satisfy the relation (21). This is a contradiction.

THEOREM X. *If all solutions satisfy* (2) [(3)] *when* $\lambda = \bar{\lambda}$, *then* $\bar{\lambda} = \lambda_j$, *where* $j$ *is determined by the number of nodes of any particular solution on any interval of length* $\omega$.

This theorem is immediate.

THEOREM XI. *If* $\lambda_j(x)$ *is constant over any interval then it is constant for all values of* $x$; *and when* $\lambda = \lambda_j$ *all solutions satisfy* (2) [(3)].

Suppose $\lambda_j(x)$ constant over some interval $x' \leqslant x \leqslant x''$. Then by Theorem IX, if $a$ belongs to this interval, $\lambda_j(a)$ is a value $l$ or $\bar{l}$ for which $\rho = 1$ or $-1$. Take two distinct points $a$ and $b$ of the interval. Then choose $y_a(x)$ and $y_b(x)$ such that $y_a(a) = 0$ and $y_b(b) = 0$ without either solution being identically zero. Then $y_a(x)$ and $y_b(x)$ are linearly independent if $a$ and $b$ are sufficiently close together. But $y_a(x)$ and $y_b(x)$ satisfy identically (2) [(3)]. Consequently, all solutions satisfy (2) [(3)] and $\lambda_j(x)$ is a constant.

THEOREM XII. *When* $\lambda_j(x)$ *is not a constant all its maxima are equal among themselves and all its minima are equal among themselves.*

Let us suppose that $\lambda_j(x)$ has two minima $l_j^{(1)}$ and $l_j^{(2)}$ and that $l_j^{(2)} < l_j^{(1)}$. Suppose that $\lambda_j(c)$ is not an extreme. Suppose, moreover, that $\lambda_j(x)$ is an increasing function at $c$. Let $x$ decrease from $c$. Then $\lambda_j(x)$ can never get less than $l_j^{(1)}$, for as soon as it equals $l_j^{(1)}$, suppose at $b$, there exists a solution, not identically zero, satisfying (2) [(3)] which we denote by $y_n$. The corresponding value of $\lambda_j(x)$, namely $\lambda_j(b) = l_j^{(1)}$, is then a minimum and the function begins to increase again as $x$ proceeds through $b$. So every time $\lambda_j(x)$ reaches $l_j^{(1)}$ it begins to increase again and hence it can never reach $l_j^{(2)}$. Similar reasoning applies to the maxima.

THEOREM XIII. *If* $\lambda_j(x)$ *is not a constant,* $j > 0$, *then it always has on every closed interval of length* $\omega$ *at least one maximum and at least one minimum.*

This theorem is immediate from the fact that $\lambda_j(x)$ has the period $\omega$.

Our next undertaking is to distinguish the two cases when the two roots of the characteristic equation are equal, namely to determine whether $\rho = 1$ or $\rho = -1$.

The following theorem is immediate, if we recall that the nodes of linearly independent solutions separate each other.

THEOREM XIV. *If $j$ is even $\rho = -1$, if $j$ is odd $\rho = 1$.*

## 7. Solutions without zeros and solutions with the maximum number of zeros

We assume $K(i)$ independent of $\lambda$, an assumption not previously made.

We have seen that maxima and minima of $\lambda_0(x)$ are values for which a solution not identically zero satisfying (3) exists. We propose the question: Are there values $l$ for which a solution, not identically zero, satisfying (2) exists, with $l > \lambda_0(x)$? We shall answer this question in the affirmative.

Let $y_1(x)$ and $y_2(x)$ be two solutions such that

$$y_1(a) = 0, \qquad y_1(a+1) = \frac{1}{K(a)};$$
$$y_2(a) = \frac{1}{K(a)}, \qquad y_2(a+1) = 0.$$

Now use these two solutions to construct the characteristic equation. We find

$$\rho^2-(\alpha_{11}+\alpha_{22})\rho+1 = 0,$$

where $\quad \alpha_{11}+\alpha_{22} = K(a)[y_1(a+\omega+1)+y_2(a+\omega+1)].$

Now when $\lambda$ is very large $|y_1(a+\omega+1)/y_2(a+\omega+1)|$ is as large as we like. This is true since $y_1(a+\omega+1)$ is of the order of $G(a, \lambda)G(a+1, \lambda)...G(a+\omega-1, \lambda)$ and $y_2(a+\omega+1)$ is of the order of $G(a+1, \lambda)...G(a+\omega-1, \lambda)$. Moreover, $K > 0$, and for large values of $\lambda$, $G > 0$. Consequently, $y_1(a+\omega+1) > 0$ since $y_1(a) = 0$ and $y_1(a+1) > 0$. Hence, if $\lambda$ is sufficiently large $\rho+\rho_2 = \alpha_{11}+\alpha_{22} > 0$ is large and positive. Moreover, $\rho_1 \rho_2 = 1$. Suppose $\rho_1$ large and $\rho_2$ small. Now $G$ and $K$ are continuous in $\lambda$ and, hence, $y_1(a+\omega+1)$ and $y_2(a+\omega+1)$ and, hence, $\rho_1$ and $\rho_2$.

Consequently, as $\lambda$ decreases, $\rho_1 = \rho_2 = 1$ before $\rho_1 = \rho_2 = -1$. In other words, a value $l$ exists larger than any value $\bar{l}$, in particular larger than the maximum of $\lambda_0(x)$.

We now propose the question: Is there more than one such value? Let us assume that there are two such values $l_0$ and $L_0$. Denote corresponding solutions satisfying (2) by $y$ and $Y$. Assume both of the same sign. If necessary, multiply one of them by a constant to bring this about. From (14)

$$[K(a+\omega)y(a+\omega)\Delta Y(a+\omega)-K(a+\omega)Y(a+\omega)\Delta y(a+\omega)]-$$
$$-[K(a)y(a)\Delta Y(a)-K(a)Y(a)\Delta y(a)]$$
$$= \sum_{i=a}^{a+\omega-1} [G(i, L_0)-G(i, l_0)]Y(i+1)y(i+1).$$

That is,

$$\sum_{i=a}^{a+\omega-1} [G(i, L_0)-G(i, l_0)]Y(i+1)y(i+1) = 0.$$

This is impossible since $G(i, \lambda)$ is an increasing function and $Y(i+1)y(i+1) > 0$. Hence no value $L$ exists.

A precisely similar line of reasoning shows that in case $\omega-1$ is odd there exists a value $l_{\omega-1} < \lambda_{\omega-2}$ and in case $\omega-1$ is even a value $\bar{l}_{\omega-1} < \lambda_{\omega-2}$.

It is only necessary to remark that for numerically large negative values of $\lambda$ we know that $y_1(a+\omega+1)$ has the sign of $G(a, \lambda)G(a+1, \lambda)...G\{a+(\omega-1), \lambda\}$, which is positive or negative as $\omega$ is even or odd. For these large negative values of $\lambda$ the sum $\rho_1+\rho_2$ is numerically large with the sign of $y_1(a+\omega+1)$ and is continuous. As $\lambda$ increases we must have $\rho_1 = \rho_2 = -1$ before $\rho_1 = \rho_2 = 1$, or vice versa according as $\omega$ is even or odd.

## 8. Sturm's normal form subject to periodic boundary conditions

Let us be given the equation

$$\Delta[K(i)\Delta y(i)]-G(i, \lambda)y(i+1) = 0, \tag{23}$$

where $G(i, \lambda)$ is defined when

$$a \leqslant i \leqslant a+\omega-1$$

and is a continuous function of $\lambda$ increasing from $-\infty$ to $\infty$ as $\lambda$ increases; where moreover $K(i)$ is defined when

$$a \leqslant i \leqslant a+\omega, \tag{24}$$

and $$K(a) = K(a+\omega).$$

Denote by $\lambda_j$ those values of $\lambda$ for which a solution, not identically zero, vanishing at $a$ vanishes also at $a+\omega$ with exactly $j$ nodes on the interval

$$a+1 < x < a+\omega-1. \tag{25}$$

We know that there exist such values as follows

$$\lambda_0 > \lambda_1 > \dots > \lambda_{\omega-2}.$$

We now define the coefficients $K(i)$ and $G(i, \lambda)$ by the formulae $K(i+\omega) = K(i)$, $G(i+\omega, \lambda) = G(i, \lambda)$ for all values of $i$ for which they are not already defined. This does not affect the solutions of the equation over the interval $a \leqslant i \leqslant a+\omega+1$. From our previous discussions of this chapter we infer the following theorem:

THEOREM XV. *There exist values $l_j$ such that, when $\lambda = l_j$, a solution not identically zero exists satisfying the boundary conditions*

$$\begin{aligned} y(a) &= y(a+\omega), \\ y(a+1) &= y(a+\omega+1) \end{aligned} \tag{26}$$

*with exactly $j$ nodes on the interval*

$$a < x < a+\omega. \tag{27}$$

*Moreover*

$$l_0 > \lambda_1 > l_1 \geqslant \lambda_2 \geqslant l_2 > \lambda_3 > l_3 \geqslant \lambda_4 \geqslant \dots \gtrless l_{\omega-1}. \tag{28}$$

*Here in the last inequality the $\geqslant$ sign holds if $\omega$ is odd and the $>$ sign if $\omega$ is even.*

*There exist values $\bar{l}_j$ such that, when $\lambda = \bar{l}_j$, a solution not identically zero exists satisfying the boundary conditions*

$$\begin{aligned} y(a) &= -y(a+\omega), \\ y(a+1) &= -y(a+\omega+1) \end{aligned} \tag{29}$$

*with exactly $j$ nodes on the interval* (27).

*Moreover*

$$\bar{l}_0 \geqslant \lambda_1 \geqslant \bar{l}_1 > \lambda_2 > \bar{l}_2 \geqslant \dots \gtrless \bar{l}_{\omega-1}. \tag{30}$$

*The final values* $l_{\omega-1}$ *and* $\bar{l}_{\omega-1}$ *are maxima of* $\lambda_{\omega-1}(x)$ *as previously discussed according as* $\omega$ *is even or odd.*

There are no values $l$ or $\bar{l}$ other than the unique ones listed in (28) and (30).

We remark that the restriction that $K$ be independent of $\lambda$ was added only to establish the existence and uniqueness of $l_0$ and $\bar{l}_{\omega-1}$ or $l_{\omega-1}$, according as $\omega-1$ is odd or even.

## 9. The non-homogeneous equation

Let the given equation be

$$y(i+2)+p(i)y(i+1)+q(i)y(i) = R(i), \tag{31}$$

when $p(i)$, $q(i)$, and $R(i)$ are defined when $-\infty < i < \infty$ and each has the period $\omega$, $q(i) \neq 0$ at any point and $R(i) \not\equiv 0$.

Let the roots of the characteristic equation of

$$y(i+2)+p(i)y(i+1)+q(i)y(i) = 0 \tag{32}$$

be $\rho_1$ and $\rho_2$ which we assume for the time being are distinct.

Particular solutions of (32) are of the form $\rho_1^{i/\omega}f_1(i)$ and $\rho_2^{i/\omega}f_2(i)$, where $f_1(i)$ and $f_2(i)$ are periodic and neither identically zero. To obtain the general solution of (31) we follow the method of variation of constants.

We have a particular solution

$$y(i) = \rho_1^{i/\omega}f_1(i)\sum\frac{-\rho_2^{(i+1)/\omega}f_2(i+1)R(i)}{W(i+1)}+$$

$$+\rho_2^{i/\omega}f_2(i)\sum\frac{\rho_1^{(i+1)/\omega}f_1(i+1)R(i)}{W(i+1)}, \tag{33}$$

$$W(i) = \rho_1^{(i+1)/\omega}f_1(i+1)\rho_2^{i/\omega}f_2(i)-\rho_2^{(i+1)/\omega}f_2(i+1)\rho_1^{i/\omega}f_1(i).$$

Here the additive arbitrary constants of summation are taken as zero. Each summand is in the form of a periodic function multiplied by $\rho_j^{-(i+1)/\omega}$, $j = 1, 2$. Expand the periodic factors into trigonometric sums (Chap. XV). Distribute the sign of summation. There is first a term of the form

$$a_0\sum\rho^{-(i+1)/\omega} = \bar{a}_0\,\rho^{-(i+1)/\omega}.$$

All other terms are of the form

$$b_k \sum \rho_j^{-(i+1)/\omega} \sin \frac{2k\pi i}{\omega} \quad \text{or} \quad a_k \sum \rho_j^{-(i+1)/\omega} \cos \frac{2k\pi i}{\omega}. \tag{34}$$

$$0 < k \leqslant \omega/2, \quad j = 1, 2.$$

We treat these by summation by parts. We find

$$\sum \rho_j^{-(i+1)/\omega} \sin \frac{2k\pi i}{\omega} = \rho_j^{-(i+1)/\omega}\left[A_1 \cos \frac{2k\pi i}{\omega} + B_1 \sin \frac{2k\pi i}{\omega}\right]$$

and

$$\sum \rho_j^{-(i+1)/\omega} \cos \frac{2k\pi i}{\omega} = \rho_j^{-(i+1)/\omega}\left[A_2 \cos \frac{2k\pi i}{\omega} + B_2 \sin \frac{2k\pi i}{\omega}\right],$$

where $A_1$, $B_1$, $A_2$, $B_2$ are constants. Substitute these in (33) and we see that we have a particular solution which is periodic. Denote this by $F(i)$. For the general solution of equation (31) we must add the general solution of (32). Hence the general solution of (31) is of the form

$$C_1 \rho_1^{i/\omega} f_1(i) + C_2 \rho_2^{i/\omega} f_2(i) + F(i), \tag{35}$$

where $f_1$, $f_2$, and $F$ are periodic with period $\omega$.

We next suppose that $\rho_1 = \rho_2$. Particular solutions of the reduced equation (32) are now of the forms $\rho_1^{i/\omega} f_1(i)$ and $\rho_1^{i/\omega} f_2(i) + ki\rho_1^{i/\omega} f_1(i)$, where $f_1(i)$ and $f_2(i)$ are periodic.

To get a particular solution of the non-homogeneous equation we again follow the method of variation of constants. As previously, summing by parts we find a particular solution of the form

$$H_1(i) + iH_2(i),$$

where $H_1(i)$ and $H_2(i)$ each has the period $\omega$. To get the general solution of (31) we add the general solution of (32), obtaining

$$\rho_1^{i/\omega}[C_1 f_1(i) + C_2\{f_2(i) + kif_1(i)\}] + H_1(i) + iH_2(i). \tag{36}$$

The results of this section can be summarized in the following theorem.

THEOREM XVI. *If $\rho_1$ and $\rho_2$, the roots of the characteristic equation of* (32) *are distinct, then* (31) *has at least one solution with the period $\omega$. If $\rho_1$ and $\rho_2$ are coincident, then* (31) *has at least one solution of the form $H_1(i) + iH_2(i)$, where $H_1(i)$ and $H_2(i)$*

*each has the period* $\omega$. *The general solution is given by* (35) *or* (36) *according as* $\rho_1 \neq \rho_2$ *or* $\rho_1 = \rho_2$.

## 10. Generalizations

The results in this chapter have been generalized.† The underlying idea of this generalization alone will be given here. An interested reader can consult the paper in question.

The fundamental facts from which the material of the chapter follows are: ($a$) If $y(i)$ is a solution of the difference equation then $y(i+\omega)$ is also a solution. ($b$) A necessary and sufficient condition that a solution $y(i)$ be periodic is $y(a) = y(a+\omega)$, $y(a+1) = y(a+\omega+1)$. Result ($a$) follows from the periodic character of the coefficients. We consequently undertake to generalize the notion of periodicity.

Let
$$u_n(i) = a_{11}(i)y_n(i)+a_{12}(i)y_n(i-1),$$
$$v_n(i) = b_{11}(i)y_n(i+1)+b_{12}(i)y_n(i).$$
Now we define our coefficients so that if $y_n(i)$ is a solution of the difference equation there exists a solution $y_{\bar{n}}(i)$ such that
$$u_n(i) \equiv v_{\bar{n}}(i+\omega).$$
Such a definition proves possible and is a generalization of periodicity. The general theory of the characteristic equation follows as for the simpler periodic case.

We now generalize condition ($b$), but for details refer to p. 14 of the paper. There is a complete theory including the solution of the boundary-value problem for Sturm's normal form as expressed briefly below:
$$\alpha_{11} K(a)\Delta y(a)+\alpha_{12} y(a) = \beta_{11} K(a+\omega)\Delta y(a+\omega)+\beta_{12} y(a+\omega),$$
$$\alpha_{21} K(a)\Delta y(a)+\alpha_{22} y(a) = \beta_{21} K(a+\omega)\Delta y(a+\omega)+\beta_{22} y(a+\omega),$$
$$\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21} = \beta_{11}\beta_{22}-\beta_{12}\beta_{21}.$$
The theory of maxima and minima of functions $\lambda_j(x)$ arises as in the periodic case.

An almost identical theory exists for the differential equation and is carried through in detail in the paper. Its proof by means of the limit theory of Chapter XI would be interesting.

† Fort, T., *Amer. Journ. of Math.* **39**, 1.

## 11. Bounded and unbounded linear equations

Given
$$y(i+2)+M(i)y(i+1)+y(i) = 0, \tag{37}$$
where the function $M(i)$ is real and defined for all integral values of the argument, and satisfies the relation $M(i+\omega) = M(i)$.

*This equation is called bounded if all solutions are bounded. It is said to be unbounded in the contrary case.*

Let $y_1$ and $y_2$ be two real linearly independent solutions. Then, as $y_1(i+\omega)$ and $y_2(i+\omega)$ are also solutions,
$$\begin{aligned} y_1(i+\omega) &= a_{11}\,y_1(i)+a_{12}\,y_2(i), \\ y_2(i+\omega) &= a_{21}\,y_1(i)+a_{22}\,y_2(i), \end{aligned} \tag{38}$$
where $a_{11}$, $a_{12}$, $a_{21}$, $a_{22}$ are constants; and since the determinant
$$\begin{vmatrix} y_1(i) & y_1(i+1) \\ y_2(i) & y_2(i+1) \end{vmatrix}$$
is a constant,
$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = 1.$$
Consequently, the characteristic equation of (37),
$$\begin{vmatrix} a_{11}-\rho & a_{12} \\ a_{21} & a_{22}-\rho \end{vmatrix} = 0,$$
reduces to $\rho^2-(a_{11}+a_{22})\rho+1 = 0$, which we write
$$\rho^2-2A\rho+1 = 0. \tag{39}$$
From the first of equations (38),
$$y_1(i+\omega) = 2Ay_1(i)+a_{12}\,y_2(i)-a_{22}\,y_1(i). \tag{40}$$
Write equations (38) in the form
$$\begin{aligned} y_1(i) &= a_{11}\,y_1(i-\omega)+a_{12}\,y_2(i-\omega), \\ y_2(i) &= a_{21}\,y_1(i-\omega)+a_{22}\,y_2(i-\omega). \end{aligned} \tag{41}$$
Solve (41) for $y_1(i-\omega)$ and substitute in (40). We get
$$y_1(i+\omega)+y_1(i-\omega) = 2Ay_1(i).$$
But $A$ is real and, being a coefficient of the characteristic equation, is independent of the particular fundamental system of solutions chosen. Hence, when $y$ is any solution of (37),
$$y(i+\omega)+y(i-\omega) = 2Ay(i). \tag{42}$$

We call $A$ the characteristic constant of the difference equation (37).

Relation (42) can be written in the form

$$y\{i+(n+2)\omega\}-2Ay\{i+(n+1)\omega\}+y(i+n\omega) = 0, \qquad (43)$$

that is, $y(i+n\omega)$ satisfies the difference equation

$$u(n+2)-2Au(n+1)+u(n) = 0. \qquad (44)$$

Solve this equation subject to the initial conditions

$$u(0) = y(i), \qquad u(1) = y(i+\omega).$$

Denote the roots of the quadratic equation $\alpha^2-2A\alpha+1 = 0$ by $\alpha_1$ and $\alpha_2$:

$$\alpha_1 = A+\sqrt{(A^2-1)}, \qquad \alpha_2 = A-\sqrt{(A^2-1)}.$$

If $A^2 \neq 1$,

$$y(i+n\omega) = C_1\alpha_1^n+C_2\alpha_2^n. \qquad (45)$$

From the initial conditions we readily determine $C_1$ and $C_2$:

$$C_1 = \frac{y(i+\omega)-\{A-\sqrt{(A^2-1)}\}y(i)}{2\sqrt{(A^2-1)}},$$

$$C_2 = \frac{y(i+\omega)-\{A+\sqrt{(A^2-1)}\}y(i)}{2\sqrt{(A^2-1)}}.$$

Let

$$T_n = \frac{\{A+\sqrt{(A^2-1)}\}^n-\{A-\sqrt{(A^2-1)}\}^n}{2\sqrt{(A^2-1)}}$$

and substitute for $C_1$ and $C_2$ in (45), then

$$y(i+n\omega) = T_n y(i+\omega)-T_{n-1}y(i). \qquad (46)$$

When $A^2 = 1$ define $T_n = \lim_{A^2=1} T_n$. Then $T_n = (\pm 1)^{n+1}n$, according as $A = \pm 1$.

We immediately verify that formula (46) still gives the required solution. It reduces to

$$y(i+n\omega) = (\pm 1)^n y(i)+(\pm 1)^{n+1}[y(i+\omega)\mp y(i)]n. \qquad (47)$$

Let us assume $A^2 > 1$ and let $A = \frac{1}{2}(a+1/a)$, $a = A\pm\sqrt{(A^2-1)}$. Choose $a = A-\sqrt{(A^2-1)}$. Substitute in the formula for $T_n$ and we get

$$T_n = \frac{(1/a)^n-a^n}{(1/a)-a}.$$

This function is unbounded as $n$ becomes infinite. Moreover one readily sees that $K_1 T_n-K_2 T_{n-1}$, where $K_1$ and $K_2$ are

independent of $n$ and not both zero, is also unbounded. Consequently, from formula (46), all solutions of (37) not identically zero are unbounded.

If $A^2 < 1$, let $A = \cos\alpha$:

$$T_n = \frac{\{\cos\alpha+\sqrt{(-1)}\sin\alpha\}^n-\{\cos\alpha-\sqrt{(-1)}\sin\alpha\}^n}{2\sqrt{(-1)}\sin\alpha}$$

$$= \frac{\sin n\alpha}{\sin\alpha},$$

which is bounded. Hence, by (46), all solutions of (37) are bounded.

If $A^2 = 1$ we see from (47) that a solution $y(i)$ is unbounded if $y(i+\omega)\mp y(i) \neq 0$, and bounded in the contrary case. But $y(i+\omega)\mp y(i)$ is a solution of the given equation. Denote it by $\theta(i)$. Apply (42) to its first term, then

$$\theta(i+\omega) = y(i+2\omega)\mp y(i+\omega) = \pm y(i+\omega)-y(i) = \pm\theta(i),$$

that is, this solution is periodic with a period $\omega$ or $2\omega$ according as $A = \pm 1$. This it will appear is in accordance with previous results.

We state the above results in the form of a theorem, and have:

THEOREM XVII. *If $A^2 > 1$ all solutions of* (37), *not identically zero, are unbounded. If $A^2 < 1$ all solutions of* (37) *are bounded. If $A = 1$, there exists at least one solution, not identically zero, having the period $\omega$; all solutions not having the period $\omega$ are unbounded. If $A = -1$, there exists at least one solution, not identically zero, satisfying the relation $y(i+\omega) = -y(i)$; all solutions not having the period $2\omega$ are unbounded.*

The problem of the calculation of $A$ next presents itself.

Let $f(i)$ and $\phi(i)$ be the two solutions of (37) such that

$$f(0) = 1, \qquad \Delta f(0) = 0;$$

$$\phi(0) = 0, \qquad \Delta\phi(0) = 1.$$

From (42), $f(\omega)+f(-\omega) = 2A$. Moreover,

$$\{\Delta\phi(\omega)\}f(i+\omega)-\{\Delta f(\omega)\}\phi(i+\omega)$$

is a linear combination of $f(i+\omega)$ and $\phi(i+\omega)$, two solutions of

(37), and hence is itself a solution. Moreover, it and its first difference at 0 are equal respectively to $f(0)$ and $\Delta f(0)$; hence

$$f(i) = \{\Delta\phi(\omega)\}f(i+\omega)-\{\Delta f(\omega)\}\phi(i+\omega). \tag{48}$$

From (48) $f(-\omega) = \Delta\phi(\omega)$; and hence

$$2A = f(\omega)+\Delta\phi(\omega). \tag{49}$$

If $\omega$ is small, the calculation of $A$ from (49) is easy. We calculate $f(2), f(3),..., f(\omega)$ successively from (37), then $\phi(2)$, $\phi(3)$,..., $\phi(\omega+1)$, and substitute in (49).

If $\omega$ is large, this process is tedious, and for very large values of $\omega$ is prohibitive. In the following pages we prove a theorem by means of which it is frequently possible to determine whether $A^2 > 1$, $A^2 = 1$, or $A^2 < 1$ with a minimum of labour. We begin by writing the difference equation in the form

$$\Delta^2 y(i)+p(i)y(i+1) = 0, \tag{50}$$

where $p(i)$ replaces $M(i)+2$.

Treat $f(i)$ and $\phi(i)$ by a method of successive approximations.† Denote the successive terms in the two series by $f_0$, $-f_1$,..., $(-1)^{\omega-1}f_{\omega-1}$ and $\phi_0$, $-\phi_1$,..., $(-1)^{\omega}\phi_\omega$ respectively. Adopting the convention $\sum_{i=k}^{g} F(i) = 0$, $k > g$, then when $i \geqslant 0$

$$\phi_0(i) = i,$$

$$\phi_1(i) = \sum_{i_1=0}^{i-1}\sum_{i_2=0}^{i_1-1} p(i_2)(i_2+1),$$

. . . . . . . . . . . . . . . .

$$\phi_n(i) = \sum_{i_1=0}^{i-1}\sum_{i_2=0}^{i_1-1}\sum_{i_3=0}^{i_2}\cdots\sum_{i_{2n-1}=0}^{i_{2n-2}}\sum_{i_{2n}=0}^{i_{2n-1}-1} p(i_2)p(i_4)...p(i_{2n})(i_{2n}+1); \tag{51}$$

$$f_0(i) = 1,$$

. . . . . . . . . . . . . . .

$$f_n(i) = \sum_{i_1=0}^{i-1}\sum_{i_2=0}^{i_1-1}\sum_{i_3=0}^{i_2}\cdots\sum_{i_{2n-1}=0}^{i_{2n-2}}\sum_{i_{2n}=0}^{i_{2n-1}-1} p(i_2)p(i_4)...p(i_{2n}). \tag{52}$$

From (51)

$$\Delta\phi_n(i) = \sum_{i_2=0}^{i-1}\sum_{i_3=0}^{i_2}\sum_{i_4=0}^{i_3-1}\cdots\sum_{i_{2n}=0}^{i_{2n-1}-1} p(i_2)p(i_4)...p(i_{2n})(i_{2n}+1)$$

$$= \sum_{i_1=0}^{i-1}\sum_{i_2=0}^{i_1}\sum_{i_3=0}^{i_2-1}\cdots\sum_{i_{2n-1}=0}^{i_{2n-2}-1}\sum_{i_{2n}=0}^{i_{2n-1}} p(i_1)p(i_2)...p(i_{2n-1}). \tag{53}$$

† See Chapter XI.

Let $$2A_n = f_n(\omega)+\Delta\phi_n(\omega).$$

Clearly $$A = 1-A_1+A_2-\ldots+(-1)^\omega A_\omega.$$

Consider $$f_2(\omega) = \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1} p(i_2)\sum_{i_3=0}^{i_2}\sum_{i_4=0}^{i_3-1} p(i_4).$$

Let $\sum_{i=0}^{i-1} p(i) = P(i)$ thus defining $P(i)$; and sum by parts considering $i_2$ as variable of summation. We get

$$f_2(\omega) = \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1}\{P(i_1)-P(i_2)\}P(i_2).$$

Now apply similar summation by parts to $f_n(\omega)$ considering successively, as variables of summation, $i_2, i_4,\ldots, i_{2n-2}$. The above result is clearly general for any single summation, and we write

$$f_n(\omega) = \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1}\sum_{i_3=0}^{i_2-1}\cdots\sum_{i_{2n}=0}^{i_{2n-1}-1}\{P(i_1)-P(i_2)\}\{P(i_2)-P(i_3)\}\ldots$$
$$\ldots\{P(i_{n-1})-P(i_n)\}. \quad (54)$$

Consider next

$$\Delta\phi_2(\omega) = \sum_{i_1=0}^{\omega-1} p(i_1)\sum_{i_2=0}^{i_1}\sum_{i_3=0}^{i_2-1} p(i_3)\sum_{i_4=0}^{i_3} 1.$$

Sum by parts, considering successively $i_1$ and $i_3$ as variables of summation:

$$\Delta\phi_2(\omega) = \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1}\{P(\omega)-P(i_1)\}\{P(i_1)-P(i_2)\}.$$

In general, letting $P(\omega) = \Omega$,

$$\Delta\phi_n(\omega) = \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1}\sum_{i_3=0}^{i_2-1}\cdots\sum_{i_n=0}^{i_{n-1}-1}\{\Omega-P(i_1)\}\{P(i_1)-P(i_2)\}\ldots$$
$$\ldots\{P(i_{n-1})-P(i_n)\}. \quad (55)$$

Combining (54) and (55),

$$2A_n = \sum_{i_1=0}^{\omega-1}\sum_{i_2=0}^{i_1-1}\sum_{i_3=0}^{i_2-1}\cdots\sum_{i_n=0}^{i_{n-1}-1}\{\Omega-P(i_1)+P(i_n)\}\times$$
$$\times\{P(i_1)-P(i_2)\}\ldots\{P(i_{n-1})-P(i_n)\}. \quad (56)$$

THEOREM XVIII. *If $p(i) \geqslant 0$ at all points, $A_n \geqslant 0$, and if $A_n = 0$ then $A_{n+1} = 0$; if $A_n \neq 0$ then*

$$\frac{A_{n+1}}{A_n} < \frac{n}{n+1}\frac{A_n}{A_{n-1}}.$$

For the proof of the theorem we refer to (56). This can be written

$$2A_n = \sum \{\Omega - P(i_1)+P(i_n)\}\{P(i_1)-P(i_2)\}...\{P(i_{n-1})-P(i_n)\},$$

where $\Sigma$ denotes the sum of all products of the form expressed, the letters $i_1, i_2,..., i_n$, chosen in every possible way from the numbers $\omega-1$, $\omega-2$,..., 1, 0, subject to the restrictions

$$i_1 > i_2 > i_3 > ... > i_n.$$

If we conceive of the numbers $\omega-1$, $\omega-2$,..., 1, 0 as equally spaced points on a circle of circumference $\omega$, in the expression

$$\{\Omega - P(i_1)+P(i_n)\}\{P(i_1)-P(i_2)\}...\{P(i_{n-1})-P(i_n)\},$$

the first factor is in no manner different from any other, and (56) can be written

$$2A_n = \sum \Big[\Big\{\sum_{i=k_0}^{k_{l_1}} p(i)\Big\}\Big\{\sum_{i=k_{l_1}+1}^{k_{l_2}} p(i)\Big\}\cdots\Big\{\sum_{i=k_{l_{n-1}}+1}^{k_{l_n}} p(i)\Big\}\Big],$$

where $\Sigma$ denotes the sum of all possible products of the form expressed, $k_0, k_1,..., k_{l_n} = k_\omega$ being the numbers $\omega-1$, $\omega-2$,..., 0 taken always in the same cyclic order, namely $\omega-1$, $\omega-2$,..., 0. For brevity we write

$$2A_n = \sum {}_0D_{l_1}\, {}_{l_1}D_{l_2} \cdots {}_{l_{n-1}}D_{l_n}.$$

Then

$$4A_n^2 = [\sum {}_0D_{\lambda_1}\, {}_{\lambda_1}D_{\lambda_2} \cdots {}_{\lambda_{n-1}}D_{\lambda_n}][\sum {}_0D_{\mu_1}\, {}_{\mu_1}D_{\mu_2} \cdots {}_{\mu_{n-1}}D_{\mu_n}], \quad (57)$$

$$4A_{n-1}A_{n+1} = [\sum {}_0D_{\nu_1}\, {}_{\nu_1}D_{\nu_2} \cdots {}_{\nu_{n-2}}D_{\nu_{n-1}}][\sum {}_0D_{\rho_1}\, {}_{\rho_1}D_{\rho_2} \cdots {}_{\rho_n}D_{\rho_{n+1}}], \quad (58)$$

where, instead of using only the letter $l$, we use distinct letters $\lambda$, $\mu$, $\nu$, $\rho$.

We shall consider (57) and (58). Begin by supposing $a_1, a_2,..., a_{2n}$ numbers of the succession $\omega-1$, $\omega-2$,..., 0 and assume that among the $a$'s there are exactly $k$ distinct numbers, and that no number occurs more than twice among them. If $k \geqslant n+1$ the product $p(a_1)p(a_2)...p(a_{2n})$ will occur in the expanded right-hand members of both (57) and (58). We shall show that the ratio of its coefficient in (57) to its coefficient in (58) is greater than or equal to $(n+1)/n$.

Omitting coefficients, let $p(j_1)...p(j_{n-1})$ be a term of $A_{n-1}$ and $p(i_1)...p(i_{n+1})$ a term of $A_{n+1}$ such that

$$p(j_1)...p(j_{n-1})p(i_1)...p(i_{n+1})$$

is identical with $p(a_1)p(a_2)...p(a_{2n})$ and let $p(\bar{j}_1)...p(\bar{j}_n)$ and $p(\bar{i}_1)...p(\bar{i}_n)$ be terms of $A_n$ such that $p(\bar{j}_1)...p(\bar{j}_n)p(\bar{i}_1)...p(\bar{i}_n)$ is identical with $p(a_1)p(a_2)...p(a_{2n})$. We shall show that the ratio of the number of ways in which $\bar{j}_1,...\bar{j}_n$, $\bar{i}_1,..., \bar{i}_n$ can be chosen to the number of ways in which $j_1,..., j_{n-1}, i_1,..., i_{n+1}$ can be chosen is greater than or equal to $(n+1)/n$. These numbers are exactly the coefficients of $p(a_1),..., p(a_{2n})$ in (57) and (58) respectively.

Require, first, that $p(j_1), p(j_2),..., p(j_{n-1})$ be each the first term of one of the parentheses ${}_0D_{\nu_1}, {}_{\nu_1}D_{\nu_2},..., {}_{\nu_{n-2}}D_{\nu_{n-1}}$ and that $p(i_1)$, $p(i_2),..., p(i_{n+1})$ be each the first term of one of the parentheses ${}_0D_{\rho_1}, {}_{\rho_1}D_{\rho_2},..., {}_{\rho_n}D_{\rho_{n+1}}$. Under this requirement the number of ways in which $j_1,..., j_{n-1}, i_1,..., i_{n+1}$ can be chosen is the number of ways in which $j_1,..., j_{n-1}$ can be chosen from $a_1, a_2,..., a_{2n}$. The $2n-k$ numbers which occur twice among $a_1, a_2,..., a_{2n}$ necessarily occur among $j_1, j_2,..., j_{n-1}$. There remain $k-n-1$ of the $j$'s which can be chosen arbitrarily from the remaining $2k-2n$ numbers. Letting $k-n = N_k$, this can be done in $\frac{2N_k(2N_k-1)...(N_k+2)}{(N_k-1)!}$ ways. Similarly, the number of ways that $\bar{j}_1,..., \bar{j}_n, \bar{i}_1,..., \bar{i}_n$ can be chosen, requiring that $p(\bar{j}_1),..., p(\bar{j}_n)$ be each the first term of one of the parentheses ${}_0D_{\lambda_1},..., {}_{\lambda_{n-1}}D_{\lambda_n}$ and $p(\bar{i}_1),..., p(\bar{i}_n)$ be each the first term of one of the parentheses ${}_0D_{\mu_1},..., {}_{\mu_{n-1}}D_{\mu_n}$, is $\frac{2N_k(2N_k-1)...(N_k+1)}{N_k!}$. The second is larger in the ratio $(N_k+1)/N_k$. But $N_k \leqslant n$ and hence

$$(N_k+1)/N_k \geqslant \frac{n+1}{n}.$$

We generalize as follows: Instead of requiring that each $p$ be the first term of a parenthesis, let us require that $p(a_1)$ be the $\eta$th, $p(a_2)$ the $\theta$th,..., $p(a_{2n})$ the $\zeta$th. For convenience we shall refer to $j_1,..., j_{n-1}$ and $\bar{j}_1,..., \bar{j}_n$ as the sets $J$ and to $i_1,..., i_{n+1}$ and $\bar{i}_1,..., \bar{i}_n$ as the sets $I$. As above, those $a$'s occurring twice among $a_1,..., a_{2n}$ necessarily occur in both the sets $J$ and $I$. Consider them as fixed. We proceed as before, choosing the remainder of the sets $J$.

It may happen that the fact that $a_m$ lies in the sets $J$ (or $I$)

requires that $a_{m+\nu}$ lie in the corresponding sets $I$ (or $J$). Thus, suppose that $a_m$ is the $\rho$th term of a parenthesis and $a_{m+\nu}$ the $\bar{\rho}$th, and suppose that $\bar{\rho} > \nu$; then, if $a_m$ is one of the set $J$, in order for $p(a_{m+\nu})$ to lie in a different parenthesis from $p(a_m)$ as it must, it must necessarily be a member of the set $I$. Moreover, the fact that $a_m$ lies in the set $J$ can require that only one of the $a$'s lie in the sets $I$; for, suppose that $p(a_{m+n})$ is the $S$th term of a parenthesis and $S > \mu \geqslant \nu$, then $p(a_{m+\nu})$ and $p(a_{m+\mu})$ belong to the sets $I$ and lie in different parentheses. Hence $S \leqslant \mu-\nu$ but $S > \mu$, a contradiction.

In the way that we are choosing the sets $J$, let us suppose all $p(a)$'s that impose any restriction on others as fixed. Let this number be $L$. Then there are thereby fixed $R$ in the sets $I$, and necessarily $R \leqslant L$. The remaining $a$'s can now be distributed in sets $J$ and $I$ at pleasure. This can be done in (58) and (57) in

$$\frac{(2N_k-L-R)(2N_k-L-R-1)...(N_k-R+2)}{(N_k-L-1)!}$$

and $$\frac{(2N_k-L-R)(2N_k-L-R-1)...(N_k-R+1)}{(N_k-L)!}$$

ways respectively. The second is the larger, in the ratio $(N_k-R+1)/(N_k-L)$, which is greater than $(n+1)/n$. We thus conclude that the coefficient of $p(a_1)...p(a_{2n})$ in (57) is greater than its coefficient in (58) by a ratio greater than or equal to $(n+1)/n$.

We have considered $k = n+1$, which exhausts the terms of (58). There are in addition in (57) terms of the form

$$p(a_1)...p(a_{2n}),$$

where $k = n$; that is, terms of the form

$$\{p(a_1)...p(a_{2n})\}^2.$$

If $A_n \neq 0$ these are not all positive. All coefficients are positive, and hence we conclude

$$A_n^2 > \frac{n+1}{n} A_{n+1} A_{n-1}, \quad A_n \neq 0,$$

$$A_n^2 = A_{n+1} A_n, \qquad A_n = 0,$$

from which we immediately draw the desired conclusion.

From Theorem XVIII one readily proves† the following theorem:

THEOREM XIX. *If $p(i) \geqslant 0$ at all points,*

*when* $-A_1+A_2-A_3+...+A_{2n} < 0$, *then* $A < 1$,

*when* $2-A_1+A_2-...-A_{2n-1} > 0$, *then* $A > -1$,

*when* $2-A_1+A_2-...+A_{2n} < 0$, *then* $A < -1$,

*when* $-A_1+A_2-A_3+...-A_{2n-1} > 0$, *then* $A > 1$.

We calculate $A_1$, $A_2$,... successively and examine the inequalities of Theorem XIX.

† See Chapter XIII, Theorem II.

## EXERCISES

1. If in (23) $G(i, \lambda) \equiv \lambda g(i)+h(i)$ note that $A$ of § 11 is a function of $\lambda$. Prove: When $\lambda$ is a double root of $A(\lambda)-1$ then and only then do all solutions of (23) satisfy (26); and when $\lambda$ is a double root of $A(\lambda)+1$ then and only then do all solutions of (23) satisfy (29).

2. By means of summation by parts arrive at formulae for the calculation of $A_1$, $A_2$, $A_3$ similar to the formulae of Chapter XIII.

## SUGGESTED REFERENCES

BIRKHOFF, G. D.: 'Existence and Oscillation Theorem for a Certain Boundary-Value Problem', *Amer. Math. Soc. Trans.* **10,** 259.

FLOQUET, M. G.: 'Sur les Équations Différentielles Linéaires à Coefficients Périodiques', *Ann. Sci. de l'École N. Sup.*, 2nd series, **12,** 147.

FORT, T.: 'Periodic Solutions of Linear Difference Equations of the Second Order and the Corresponding Developments of Functions of an Integral Argument', *Quart. Jour. of Pure and Ap. Math.*, No. 181, 1.

—— 'Linear Differential and Difference Equations', *Amer. Jour. of Math.* **39,** 1.

—— 'Limited and Illimited Linear Difference Equations of the Second Order with Periodic Coefficients', ibid. **37,** 43.

LIAPOUNOFF, A.: 'Sur une Équation Différentielle Linéaire du Second Ordre', *C.R.* **128,** 910.

—— 'Sur une Série dans la Théorie des Équations Différentielles Périodiques du Second Ordre à Coefficients Périodiques', *Mém. de l'Académie Imp. de St. Pét.* 8th series, **13.**

MOULTON, F. R., and MACMILLAN, W. D.: 'On the Solution of Certain Types of Linear Differential Equations with Periodic Coefficients', *Amer. Jour. of Math.* **33,** 63.

POINCARÉ, H.: *Les Méthodes Nouvelles de la Mécanique Céleste,* **1,** chap. ii.

XV

# ORTHOGONAL SETS AND THE DEVELOPMENT OF AN ARBITRARY FUNCTION

## 1. Definition

*A set of functions* $y_1(i)$, $y_2(i),\ldots, y_m(i)$ *of the integral argument,* $i$, *defined when*

$$c \leqslant i \leqslant b, \tag{1}$$

*is said to be orthogonal by summation over* (1) *with respect to* $g(i)$ *if*

$$\sum_{i=c}^{b} g(i-1)y_j(i)y_k(i) = 0, \quad j \neq k. \tag{2}$$

We have encountered several such sets of functions.†

## 2. The set satisfying periodic boundary conditions

Consider the difference equation

$$\Delta^2 y(i) - [\lambda g(i) + h(i)]y(i+1) = 0, \tag{3}$$

where the real functions $g$ and $h$ are defined at every point of the interval

$$c-1 \leqslant i \leqslant b-1 \quad \text{and} \quad g(i) > 0. \tag{4}$$

In Chapter XIV it was proved that there exist values of $\lambda$

$$l_1 > l_2 \geqslant l_3 > l_4 \geqslant l_5 > \ldots \gtreqless l_{b-c+1}$$

such that when $\lambda = l_j$, $j = 1, 2,\ldots, b-c+1$, there exist solutions of (3), not identically zero, which satisfy the conditions

$$y(c-1) = y(b), \qquad y(c) = y(b+1); \tag{5}$$

moreover that there exist two linearly independent solutions satisfying (5) when and only when $\lambda = l_{2n} = l_{2n+1}$.

Suppose $l_{2n} > l_{2n+1}$. When $\lambda = l_{2n}$ or $l_{2n+1}$ we choose a real solution of (3), satisfying (5) and not identically zero, which we shall say corresponds to $l_{2n}$ or $l_{2n-1}$ as the case may be.

If $l_{2n} = l_{2n+1}$ we choose a pair of real linearly independent solutions, which we shall say correspond to $l_{2n}$ and $l_{2n+1}$ respectively. The solutions chosen will be further particularized later.

Let $y_1$, $y_2,\ldots, y_{b-c+1}$ respectively represent the above solutions.

† See, for example, Chapter XII, § 1.

Suppose $l_k \neq l_j$. Then, over (1),

$$\Delta^2 y_j(i-1)-[l_j g(i-1)+h(i-1)]y_j(i) = 0,$$
$$\Delta^2 y_k(i-1)-[l_k g(i-1)+h(i-1)]y_k(i) = 0.$$

Multiply the first by $y_k(i)$ and the second by $y_j(i)$ and subtract:

$$\Delta[y_k(i-1)\Delta y_j(i-1)-y_j(i-1)\Delta y_k(i-1)]$$
$$= (l_j-l_k)g(i-1)y_j(i)y_k(i).$$

Sum from $c$ to $b$:

$$[y_k(i-1)\Delta y_j(i-1)-y_j(i-1)\Delta y_k(i-1)]_c^{b+1}$$
$$= \sum_{i=c}^{b}(l_j-l_k)g(i-1)y_j(i)y_k(i).$$

But from conditions (5) the left-hand member is zero. Moreover $l_j-l_k \neq 0$. Hence

$$\sum_{i=c}^{b} g(i-1)y_j(i)y_k(i) = 0, \quad j \neq k. \tag{6}$$

Hence $y_1, y_2,\ldots, y_{b-c+1}$ constitute an orthogonal system.

If $l_{2n} = l_{2n+1}$ we choose $y_{2n}$ and $y_{2n+1}$ real and linearly independent and so that (6) holds when $k$ and $j$ are replaced by $2n$ and $2n+1$ respectively. To show that this is always possible proceed as follows: Let $\bar{y}$ and $\bar{\bar{y}}$ be a pair of real linearly independent solutions of (3), for which by hypothesis (6) does not hold. Form two solutions $y_j = c_1\bar{y}+c_2\bar{\bar{y}}$ and $y_k = d_1\bar{y}+d_2\bar{\bar{y}}$, where as yet $c_1$, $c_2$, $d_1$, $d_2$ are not specified. We shall choose them so that (6) holds and so that $y_j$ and $y_k$ are linearly independent. Substitute $y_j$ and $y_k$ in the left-hand member of (6). We have

$$\sum_{i=c}^{b} g(i-1)(c_1\bar{y}+c_2\bar{\bar{y}})(d_1\bar{y}+d_2\bar{\bar{y}})$$
$$= c_1 d_1 \sum_{i=c}^{b} g(i-1)\bar{y}^2+(c_1 d_2+c_2 d_1)\sum_{i=c}^{b} g(i-1)\bar{y}\bar{\bar{y}}+$$
$$+c_2 d_2 \sum_{i=c}^{b} g(i-1)\bar{\bar{y}}^2.$$

We write this for brevity

$$A\sum_{i=c}^{b} g(i-1)\bar{y}^2+B\sum_{i=c}^{b} g(i-1)\bar{y}\bar{\bar{y}}+C\sum_{i=c}^{b} g(i-1)\bar{\bar{y}}^2 \equiv Ar+Bs+Ct.$$

Here $r > 0$, $t > 0$, $s \neq 0$.

Since, by hypothesis,

$$\sum_{i=c}^{b} g(i-1)\bar{y}\bar{\bar{y}} \neq 0$$

we can choose two numbers $A$ and $C$ at pleasure and determine a third number $B$, so that

$$Ar+Bs+Ct = 0. \tag{7}$$

We begin by requiring that $C \neq 0$. The equations $c_1 d_1 = A$, $c_1 d_2 + c_2 d_1 = B$, and $c_2 d_2 = C$, serve to determine the ratios $c_1/c_2$ and $d_1/d_2$ as two different real numbers if $A$, $B$, and $C$ are real and $B^2-4AC > 0$. Substitute for $B$ in $B^2-4AC \leqslant 0$ from (7). There results the condition

$$r^2A^2+2tCrA+t^2C^2-4ACs^2 \leqslant 0. \tag{8}$$

It is possible to choose $A$ and $C \neq 0$ as real numbers, so that (8) is not satisfied. Let them be so chosen. $B$ is then also real. It results that the corresponding values of $c_1/c_2$ and $d_1/d_2$ are real and different. Now in addition choose $c_1$, $c_2$, $d_1$, and $d_2$ real. With this choice $y_j$ and $y_k$ are real and linearly independent. They satisfy (6).

THEOREM I. *In* $y_1, y_2,\ldots, y_{b-c+1}$ *we have* $b-c+1$ *functions which are linearly independent over* (1).

Assume the contrary, that is, that

$$c_1 y_1 + c_2 y_2 + \ldots + c_{b-c+1} y_{b-c+1} = 0$$

at all points of (1) and $c_j \neq 0$. Multiply by $g(i-1)y_j(i)$, and sum from $c$ to $b$,

$$c_j \sum_{i=c}^{b} g(i-1)[y_j(i)]^2 = 0.$$

But $g(i) > 0$; and $y_j(i)$ is real and not identically zero. Hence $c_j = 0$, a contradiction.

THEOREM II. *An arbitrary function* $\phi(i)$, *defined at all points of* (1), *can be written in one and only one way as a linear function of* $y_1, y_2,\ldots, y_{b-c+1}$ *with constant coefficients.*

By means of (6) we can determine the coefficients in this development. Write

$$\phi(i) = a_1 y_1(i) + a_2 y_2(i) + \ldots + a_{b-c+1} y_{b-c+1}(i).$$

Multiply by $g(i-1)y_j(i)$, and sum from $c$ to $b$,

$$\sum_{i=c}^{b} \phi(i)g(i-1)y_j(i) = a_j \sum_{i=c}^{b} g(i-1)[y_j(i)]^2,$$

$$a_j = \frac{\sum_{i=c}^{b} \phi(i)g(i-1)y_j(i)}{\sum_{i=c}^{b} g(i-1)[y_j(i)]^2}. \tag{9}$$

The functions $y_j(i)$ can be multiplied by real constants different from zero at pleasure. If we desire we can choose constant multipliers so that the denominator in (9) is unity.

## 3. A trigonometric development

As a special case under the preceding section we obtain a well-known trigonometric development.

Consider
$$\Delta^2 y(i)+\lambda y(i+1) = 0. \tag{10}$$
Let $c = 0$ and $b = \omega-1$. When $\lambda = 0$ the constant 1 is a solution of (10) satisfying (5). If $\omega$ is even let $\omega = 2k+2$, and if it is odd let $\omega = 2k+1$. Then, when

$$\lambda = 2\left(1-\cos n\frac{2\pi}{\omega}\right), \quad 1 \leqslant n \leqslant k,$$

$\cos(2n\pi i/\omega)$ and $\sin(2n\pi i/\omega)$ are solutions of (10), which satisfy (5). If $\omega$ is even; when $\lambda = 4$, then $\cos \pi i$ is also a solution of (10) satisfying (5). Now remark that if $p \geqslant 0$ and $q \geqslant 0$ are integers, not both zero, and if $p+q < 2\omega$ is even,

$$\begin{aligned}
&\sum_{i=0}^{\omega-1} \cos p\frac{\pi i}{\omega} \cos q\frac{\pi i}{\omega} = \begin{cases} 0, & p \neq q, \\ \tfrac{1}{2}\omega, & p = q, \end{cases} \\
&\sum_{i=0}^{\omega-1} \sin p\frac{\pi i}{\omega} \cos q\frac{\pi i}{\omega} = 0, \\
&\sum_{i=0}^{\omega-1} \sin p\frac{\pi i}{\omega} \sin q\frac{\pi i}{\omega} = \begin{cases} 0, & p \neq q, \\ \tfrac{1}{2}\omega, & p = q. \end{cases}
\end{aligned} \tag{11}$$

From these formulae the above solutions satisfy (6), and hence, if $\omega = 2k+1$,

$$\phi(i) = \tfrac{1}{2}a_0 + \sum_{n=1}^{k}\left(a_n \cos n\frac{2\pi i}{\omega} + b_n \sin n\frac{2\pi i}{\omega}\right), \tag{12}$$

and if $\omega = 2k+2$,

$$\phi(i) = \tfrac{1}{2}a_0 + \sum_{n=1}^{k}\left(a_n \cos n\frac{2\pi i}{\omega} + b_n \sin n\frac{2\pi i}{\omega}\right) + \tfrac{1}{2}a_{k+1}\cos \pi i. \quad (13)$$

From (9) and (11),

$$a_n = \frac{2}{\omega}\sum_{i=0}^{\omega-1}\phi(i)\cos n\frac{2\pi i}{\omega},$$

$$b_n = \frac{2}{\omega}\sum_{i=0}^{\omega-1}\phi(i)\sin n\frac{2\pi i}{\omega},$$

$$n = 0, 1,..., k+1.$$

## 4. The set satisfying anti-periodic boundary conditions

In direct line with § 2 is the following.

Consider again (3). There exist values of $\lambda$

$$\bar{l}_1 \geqslant \bar{l}_2 > \bar{l}_3 \geqslant \bar{l}_4 > ... \gtreqless \bar{l}_{b-c+1}$$

for which there are solutions of (3), not identically zero, satisfying the conditions

$$\bar{y}(c-1) = -\bar{y}(b), \qquad \bar{y}(c) = -\bar{y}(b+1). \quad (14)$$

Proceed just as in § 2, and we find that if $\bar{l}_j \neq \bar{l}_k$ corresponding solutions satisfy the relation

$$\sum_{i=c}^{b} g(i-1)\bar{y}_j(i)\bar{y}_k(i) = 0 \quad (15)$$

and that if two $\bar{l}$'s coincide, linearly independent solutions can always be chosen satisfying this relation. This granted, like the $y$'s of § 2, the solutions corresponding to the $\bar{l}$'s, which we designate by $\bar{y}_1, \bar{y}_2,..., \bar{y}_{b-c+1}$, are linearly independent over (1), and hence we have the theorem:

THEOREM III. *If $\phi(i)$ is an arbitrary function defined at all points of* (1),

$$\phi(i) = \sum_{n=1}^{b-c+1} a_n \bar{y}_n(i),$$

*where the a's are constants. There is only one such development.*

The coefficients are determined precisely as in § 2:

$$a_n = \frac{\sum\limits_{i=c}^{b} \phi(i)g(i-1)\bar{y}_n(i)}{\sum\limits_{i=c}^{b} g(i-1)[\bar{y}_n(i)]^2}. \tag{16}$$

## 5. A trigonometric development

Again, as a special case under the preceding section we obtain a trigonometric development.

Consider (10) and let $c = 0$ and $b = \omega-1$.

Suppose $\omega$ even, $\omega = 2k$. Let $\lambda = 2\{1-\cos n(\pi/\omega)\}$, $n = 1, 3, 5,..., \omega-1$. Then $\cos n(\pi i/\omega)$ and $\sin n(\pi i/\omega)$ give $\omega$ solutions, not identically zero, which satisfy (14). They also satisfy (15), and consequently

$$\phi(i) = \sum_{n=0}^{k-1} \left[a_{2n+1} \cos(2n+1)\frac{\pi i}{\omega} + b_{2n+1} \sin(2n+1)\frac{\pi i}{\omega}\right]. \tag{17}$$

From (11) and (16),

$$a_n = \frac{2}{\omega} \sum_{i=0}^{\omega-1} \phi(i) \cos n\frac{\pi i}{\omega},$$

$$b_n = \frac{2}{\omega} \sum_{i=0}^{\omega-1} \phi(i) \sin n\frac{\pi i}{\omega}. \tag{18}$$

If $\omega$ is odd, $\omega = 2k+1$, we again let $\lambda = 2\{1-\cos n(\pi/\omega)\}$, where now $n = 1, 3,..., \omega$. We obtain the development

$$\phi(i) = \sum_{n=0}^{k-1} \left\{a_{2n+1} \cos(2n+1)\frac{\pi i}{\omega} + b_{2n+1} \sin(2n+1)\frac{\pi i}{\omega}\right\} + \tfrac{1}{2}a_\omega \cos \pi i. \tag{19}$$

The coefficients are again given by (18).

### EXERCISES

1. Consider (3) subject to the boundary conditions

$$y(c-1) = y(b+1) = 0.$$

This is a Sturm-Liouville problem discussed in Chapter X.

Prove that the characteristic functions form an orthogonal set of linearly independent functions.

2. Discuss a trigonometric development following from the discussion under Exercise 1.

3. Generalize equation (3) but do not thus destroy the orthogonal character of the sets of functions corresponding to the orthogonal sets of the text.

4. Obtain a trigonometric development for $i^2$ over $0 \leqslant i \leqslant 100$.

## SUGGESTED REFERENCE

FORT, T.: 'Periodic Solutions of Linear Difference Equations of the Second Order and the Corresponding Developments of Functions of an Integral Argument', *Quart. Jour. of Pure and Ap. Math.*, No. 181, 1.

# XVI

# OSCILLATORY AND NON-OSCILLATORY LINEAR DIFFERENCE EQUATIONS OF THE SECOND ORDER

THE present chapter is a continuation of our study of the linear recurrent relation. Some of the results are closely related to the results of Chapter X and some closely related to the results given in Chapter XIV.

## 1. Discussion for a finite interval

Given
$$y(i+2)+P(i)y(i+1)+Q(i)y(i) = 0, \tag{1}$$
where $P(i)$ and $Q(i) > 0$ are real and defined,
$$a \leqslant i \leqslant b-1. \tag{2}$$
*This equation is said to be oscillatory or non-oscillatory over the interval*
$$a \leqslant x \leqslant b \tag{3}$$
*according as it does or does not have at least one solution, not identically zero, which has more than one node on this interval.*

THEOREM I. *A necessary and sufficient condition that* (1) *be non-oscillatory over* (3) *is that it have a solution not identically zero, with no node on this interval.*

That this is sufficient follows from the fact that the nodes of two linearly independent solutions of (1) separate each other.

It is also necessary; consider two solutions $y_1(i)$ and $y_2(i)$ determined by the conditions
$$y_1(a) = 0, \qquad y_1(a+1) = 1;$$
$$y_2(b) = 0, \qquad y_2(b-1) = 1.$$
These must each be positive throughout the whole interval, the points $a$ and $b$ respectively excepted. It follows that $y_1(i)+y_2(i)$, which is also a solution, is positive without exception over (3).

COROLLARY. *A necessary and sufficient condition that* (1) *be non-oscillatory over* (3) *is that a solution, zero at a and not identically zero, have no additional node on this interval.*

Consider two equations

$$\Delta[K_1(i)\Delta y(i)]-G_1(i)y(i+1) = 0 \tag{4}$$

and

$$\Delta[K_2(i)\Delta y(i)]-G_2(i)y(i+1) = 0, \tag{5}$$

where $K_1(i)$, $K_2(i)$, $G_1(i)$, and $G_2(i)$ are defined over (2) and where $K_2(i) \geqslant K_1(i)$ and $G_2(i) \geqslant G_1(i)$.

THEOREM II. *If* (4) *is non-oscillatory over* (3), *then* (5) *is also non-oscillatory over* (3). *If* (5) *is oscillatory over* (3), *then* (4) *is also oscillatory over* (3).

Suppose (4) non-oscillatory over (3) and consider a solution satisfying the conditions

$$y(a) = 0, \qquad y(a+1) = 1. \tag{6}$$

From Theorem I corollary, this has no other node on (3). But from Chapter X, Theorem II, this solution† has at least as many nodes on (3) as a solution of (5), satisfying the same conditions at $a$. Hence, from Theorem I corollary, (5) is non-oscillatory over (3).

Suppose (5) oscillatory over (3), then from Theorem I corollary a solution satisfying the conditions $y(a) = 0$, $y(a+1) = 1$ must have at least two nodes on (3). Then a solution of (4) satisfying these same conditions at $a$ must have at least two nodes on (3) and hence, by definition, (4) is oscillatory.

THEOREM III. *A necessary and sufficient condition that* (4) *be non-oscillatory over* (3) *is that there exist a function* $\phi(i) > 0$ *over*

$$a < i \leqslant b, \tag{7}$$

$\phi(a) \geqslant 0$, *and such that*

$$\Delta[K_1(i)\Delta\phi(i)]-G_1(i)\phi(i+1) \geqslant 0 \tag{8}$$

*over* (2).

That this condition is necessary is immediate. By Theorem I there exists a solution $\phi(i) > 0$ over $a \leqslant i \leqslant b$ and hence

$$\Delta[K_1(i)\Delta\phi(i)]-G_1(i)\phi(i+1) = 0$$

over (2).

In order to prove the condition sufficient assume that there

† Theorem II of Chapter X will still hold if the assumption $G_2(i) > G_1(i)$ is replaced by $G_2(i) \geqslant G_1(i)$.

exists a $\phi(i) > 0$ over (2) satisfying (8). If necessary decrease $G_1$ to $\bar{G}_2$, so that $\Delta[K_1(i)\Delta\phi(i)]-\bar{G}_2(i)\phi(i+1) = 0$ over (2). Consider this as an equation in $\phi$. By hypothesis it has a solution $\phi(i) > 0$ over (7), and hence, by Theorem I or Theorem I corollary, it is non-oscillatory over (3). Then from Theorem II, (4) is non-oscillatory over (3).

Equation (1) can be written in the form (4), hence the above theorem can be stated in the following form:

THEOREM IV. *A necessary and sufficient condition that* (1) *be non-oscillatory over* (3) *is that there exist a function* $\phi(i) > 0$ *over* (7), $\phi(a) \geqslant 0$, *and such that*

$$\phi(i+2)+P(i)\phi(i+1)+Q(i)\phi(i) \geqslant 0 \tag{9}$$

*over* (2).

Functions $\phi(i)$ can be assumed at pleasure and resulting sufficient conditions that (1) be non-oscillatory over (3) obtained. For example, assume $\phi \equiv 1$, and we obtain the condition $G(i) \leqslant 0$; assume $\phi(i) = \alpha^i$, where $\alpha > 0$, $\alpha \neq 1$, and we obtain

$$(\alpha-1)[\alpha K(i+1)-K(i)]-G(i)\alpha \geqslant 0, \quad \text{etc.}$$

THEOREM V. *If* $r(i)$ *is a real function, which throughout* (2) *satisfies the relation* $r(i) \geqslant 0$, *then a necessary and sufficient condition that* (1) *be non-oscillatory over* (3) *is that the non-homogeneous equation*

$$\Delta[K(i)\Delta y(i)]-G(i)y(i+1) = r(i) \tag{10}$$

*have a solution satisfying the relations* $y(i) > 0$ *over* (7), $y(a) \geqslant 0$.

That this is sufficient is immediate from Theorem IV. It is also necessary, as is seen from the fact that if (1) is non-oscillatory over (3), it has a solution positive at all points of (3). By adding a sufficiently large positive multiple of such a solution of (1) to an arbitrary solution of (10), we get a solution of (10) of the form desired.

## 2. Discussion for the infinite interval

In the preceding discussion we have limited ourselves to the consideration of a finite interval. There is nothing, however, in this that is essential, and the interval $a \leqslant i \leqslant b-1$ can be

replaced by $a \leqslant i < \infty$ without in any way affecting the theorems.

THEOREM VI. *Given*

$$\Delta^2 y(i) + p(i)y(i+1) = 0,$$

*where* $p(i) > \epsilon > 0$ *is defined,*

$$a \leqslant i < \infty, \tag{11}$$

*then every solution, not identically zero, has an infinite number of nodes on the interval*

$$a \leqslant x < \infty. \tag{11'}$$

Replace $p$ by $\epsilon$. The equation $\Delta^2 y(i) + \epsilon y(i+1) = 0$ can be solved explicitly. Particular solutions are

$$\left(\frac{2-\epsilon+\sqrt{\{(2-\epsilon)^2-4\}}}{2}\right)^i \quad \text{and} \quad \left(\frac{2-\epsilon-\sqrt{\{(2-\epsilon)^2-4\}}}{2}\right)^i.$$

If $\epsilon \geqslant 4$ these solutions are real and each has an infinite number of nodes on (11′). If $\epsilon < 4$ the above solutions are imaginary, but particular real solutions are given by $r^i \cos i\theta$ and $r^i \sin i\theta$, where

$$\tfrac{1}{2}[2-\epsilon+\sqrt{\{(2-\epsilon)^2-4\}}] = r\{\cos\theta + \sqrt{(-1)}\sin\theta\}, \quad -\pi < \theta < \pi.$$

Each of these solutions has an infinite number of nodes on (11′). Moreover, since the nodes of linearly independent solutions separate each other and of linearly dependent solutions coincide, every solution has an infinite number of nodes on (11′). Now apply Chapter X, Theorem II, and the proof is complete.

In view of the above and subsequent theorems, when dealing with the infinite interval, it is convenient to make new definitions of oscillatory and non-oscillatory equations.

*An equation*

$$\Delta^2 y(i) + p(i)\Delta y(i) + q(i)y(i) = 0,$$

*where $p$ and $q$ are real and defined over* (11) *and where*

$$q(i) - p(i) > -1,$$

*will be said to be non-oscillatory if there exists a point $x_0$ beyond which no solution has more than one node. It will be said to be oscillatory in the contrary case.*

Similarly: *A differential equation*

$$\frac{d^2y}{dx^2}+\bar{p}(x)\frac{dy}{dx}+\bar{q}(x)y = 0,$$

*where the real functions $\bar{p}$ and $\bar{q}$ are defined and continuous over* (11′), *will be said to be non-oscillatory or oscillatory according as there does or does not exist a point $X_0$ beyond which no solution, not identically zero, has more than one zero.*

Consider the equation

$$\Delta^2 y(i)+p(i,\, l)\Delta y(i)+q(i,\, l)y(i+1) = 0, \tag{12}$$

where $p(x, l)$ and $q(x, l)$, as functions of the continuous variable $x$, are defined over (11′) and where $lp(lx, l)$ and $l^2q(lx, l)$ approach uniformly continuous functions $P(x)$ and $Q(x)$ respectively when $l$ becomes infinite and where $1-p(x, l) > 0$.

Assume $a > 0$ and an integer. This result can always be attained by a change of variable $j = i+C$, where $C$ is a properly chosen constant.

In (12) make the change of independent variable $i = lx$, where $l > 0$ is an integer. The integral points 0, 1, 2, 3,... are transformed into the set of points 0, $1/l$, $2/l$, $3/l$,... . Denote the abscissae of those points $1/l$, $2/l$,... which lie to the right of $a$ by $x_1$, $x_2$,... respectively.

Denote $y(lx_i)$ by $V^{(l)}(x_i)$, and (12) goes into

$$\Delta^2 V^{(l)}(x_i)+p(lx_i,\, l)\Delta V^{(l)}(x_i)+q(lx_i,\, l)V^{(l)}(x_i) = 0. \tag{13}$$

Divide through by $1/l^2 = (\Delta x_i)^2$ and we have

$$l^2\Delta^2 V^{(l)}(x_i)+lp(lx_i,\, l)l\Delta V^{(l)}(x_i)+l^2q(lx_i,\, l)V^{(l)}(x_i) = 0. \tag{14}$$

We will consider side by side with equation (14) the equation

$$y''+P(x)y'+Q(x)y = 0, \tag{15}$$

where, as usual, accents denote differentiation.

THEOREM VII. *Let $V^{(l)}(x_i)$ be a solution of* (14) *such that $V^{(l)}(a) = k$, $l\Delta V^{(l)}(a) = C$, and let $y(x)$ be a solution of* (15) *such that $y(a) = k$, $y'(a) = C$. Let $\eta > 0$ be arbitrarily small and let*

*$x_0 > a$ be a fixed point. Then we can choose a positive integer $T$ such that, when $l \geqslant T$,*

$$|V^{(l)}(x_i)-y(x_i)| < \eta,$$
$$|l\Delta V^{(l)}(x_i)-y'(x_i)| < \eta$$

*for all points $x_i$ of $a \leqslant x \leqslant x_0$.*

We proved this theorem in Chapter XI.

We can define $V^{(l)}(x)$ as a continuous function of $x$, the definition being that function defined by the broken-line graph of $V^{(l)}(x_i)$.

THEOREM VIII. *Given an $\epsilon > 0$, there exists a positive integer $T'$ such that, when $a \leqslant x \leqslant x_0$ and $l \geqslant T'$,*

$$|y(x)-V^{(l)}(x)| < \epsilon,$$
$$|y'(x)-l\Delta V^{(l)}(x)| < \epsilon.$$

Let $\eta$ of the previous theorem be less than $\frac{1}{2}\epsilon$.

$y(x)$ and $y'(x)$ are uniformly continuous over the interval $a \leqslant x \leqslant x_0$. Consequently we can choose $T' \geqslant T$ of the previous theorem so that when $l \geqslant T'$ and $a \leqslant x' \leqslant x'' \leqslant x_0$ and $x''-x' \leqslant 1/l$, $|y(x')-y(x'')| < \frac{1}{2}\epsilon$ and $|y'(x')-y'(x'')| < \frac{1}{2}\epsilon$. This $T'$ can be immediately shown to be as required in the theorem.

THEOREM IX. *A necessary and sufficient condition that* (12) *be oscillatory is that* (15) *be oscillatory.*

The oscillatory or non-oscillatory character of (12) is clearly invariant under the transformation previously made ($i = lx_i$) and hence, instead of (12), we shall consider (14).

The differential equation (15) is independent of $l$. Assume that it is oscillatory and suppose that the difference equation (14) is non-oscillatory. Let $c$ be a point beyond which no solution of (14) has more than one node when $l = l_1$. As $l$ increases, the nodes of every solution of (14), with fixed values at $a$ and $a+1/l$, move towards the left; hence, when $l \geqslant l_1$, $V^l(x_i)$ has not more than one node to the right of $c$. Take a point $b$ such that $b-c > 0$, then take an interval $b \leqslant x \leqslant d$ in which $y$ has two zeros. Denote these zeros by $S_1$ and $S_2$. As $y$

and $y'$ do not vanish together we can take a positive $\delta < \zeta$ so that in each of the intervals

$$S_1-\delta \leqslant x \leqslant S_1+\delta \quad \text{and} \quad S_2-\delta \leqslant x \leqslant S_2+\delta$$

there is only one zero of $y$. Then $y(S_1-\delta)$ and $y(S_1+\delta)$ are of opposite signs. Similarly $y(S_2-\delta)$ and $y(S_2+\delta)$ are of opposite signs. Suppose $|y(S_1 \pm \delta)| > \epsilon > 0$ and $|y(S_2 \pm \delta)| > \epsilon$. Now choose $T$ so large that, when $l \geqslant T$, $|V^{(l)}_{(x)}-y(x)| < \epsilon$ over $b \leqslant x \leqslant d$. When $l \geqslant T$, $V^{(l)}(x)$ has at least one node in each of the intervals $S_1-\delta \leqslant x \leqslant S_1+\delta$ and $S_2-\delta \leqslant x \leqslant S_2+\delta$. This is a contradiction. Hence (14) is oscillatory.

Now assume (15) non-oscillatory, but (14) oscillatory. Let $c$ be a point at which $y \neq 0$ and beyond which $y$ has no roots. Take an arbitrary interval $c \leqslant x \leqslant b$. There exists an $\epsilon > 0$ such that $|y(x)| > \epsilon$ over $c \leqslant x \leqslant b$. Let $M$ be so chosen that, when $l \geqslant M$, $|V^{(l)}(x)-y(x)| < \epsilon$ over the interval $c \leqslant x \leqslant b$.

Consider next a certain value of $l$, $l_1 > M$, and suppose that $V^{(l_1)}(x)$ has at least two nodes on the interval $c \leqslant x \leqslant d$, where $d > b$. These must lie to the right of $b$. Denote them by $S_1$ and $S_2$ ($S_1 < S_2$). We now let $l$ increase through integral values. The points $S_1$ and $S_2$ will be considered as variable, and we write $S_1(l)$ and $S_2(l)$. As $l$ increases, $S_1(l)$ and $S_2(l)$ move towards the origin, always remaining nodes of a solution of (14). Hence each subsequent $V^{(l)}$ must have nodes coinciding with $S_1(l)$ and $S_2(l)$, or at least one node lying between them. Every time $l$ is increased by unity, $S_1(l)$ and $S_2(l)$ move towards the origin distances $S_1(l)/(l+1)$ and $S_2(l)/(l+1)$ respectively. Hence

$$S_2(l+1)-S_1(l+1) < S_2(l)-S_1(l).$$

The point $b$ was arbitrary in the beginning, therefore assume that

$$b-c > S_2(l_1)-S_1(l_1)+\frac{S_2(l_1)}{l_1+1}.$$

It results that, as $l$ increases through integral values, $S_1(l)$ and $S_2(l)$ cannot get to the left of $c$ unless at some time both are on the interval $c \leqslant x \leqslant b$ at the same time. If this occurred it would necessitate a node of $V^{(l)}(x)$ on the interval $c \leqslant x \leqslant b$, which is impossible. Hence, when $l \geqslant l_1$, $S_1(l)$ and $S_2(l)$ lie on

the interval $b < x \leqslant d$ and consequently $V^{(l)}(x)$ always has at least one node on the interval $b < x \leqslant d$.

Suppose $|y(x)| > \epsilon$ over the interval $c \leqslant x \leqslant d$. We can choose $l'$ so that, when $l \geqslant l'$, $|V^{(l)}(x)-y(x)| < \epsilon$ over the interval $c \leqslant x \leqslant d$; and hence, when $l \geqslant l'$, $V^{(l)}(x)$ has no node on the interval $c \leqslant x \leqslant d$. This is a contradiction, and the theorem is complete.

It is possible to choose $\phi(i) > 0$ so that, by the change of variable $y(i) = \phi(i)\bar{y}(i)$, equation (12) is transformed into an equation of the form

$$\Delta^2 y(i)+\bar{p}(i, l)y(i+1) = 0. \tag{16}$$

As $\phi(i) > 0$ the oscillatory or non-oscillatory character of equation (12) is invariant under this transformation. The discussion can consequently be confined to (16).

When $l^2\bar{p}(lx, l)$ approaches uniformly a limit, $\bar{P}(x)$, as $l$ becomes infinite, the question of the oscillatory or non-oscillatory character of (16) is, by Theorem IX, identical with the question of the oscillatory or non-oscillatory character of the equation

$$\frac{d^2y}{dx^2}+\bar{P}(x)y = 0. \tag{17}$$

## 3. The difference equation with periodic coefficients in the infinite interval

Let $p(i+\omega, l) \equiv p(i, l)$ and let $\rho_1$ and $\rho_2$ be the roots of the characteristic equation.

Assume that $\rho_1$ and $\rho_2$ are real. There consequently exists a real solution, not identically zero, which is periodic of the second kind. Denote one such by $y_1$.

THEOREM X. *A necessary and sufficient condition that* (16) *be oscillatory is that $y_1$ have a node.*

Suppose a node of $y_1$ to lie at a point $x_0$, then nodes lie at $x_0+n\omega$, where $n = \pm 1, \pm 2, \pm 3,\ldots$. Hence the sufficiency of the condition.

If (16) is oscillatory, every solution not identically zero has an infinite number of nodes; hence the necessity of the condition.

Now $y_1$ has the same number of nodes on each interval $a+(n-1)\omega \leqslant x \leqslant a+n\omega$. Denote this number by $M$. The number of nodes on an interval $a \leqslant x < n\omega$ is then equal to $nM$. Consider an arbitrary interval $a \leqslant x < b$. Choose $k$ so that

$$0 \leqslant b-a-k\omega < \omega;$$

that is, $b-a = k\omega+m$, where $m < \omega$. The number of nodes of $y_1$ on $a \leqslant x < b$ equals $kM+K$, where $0 \leqslant K \leqslant M$. Moreover,

$$\lim_{k\to\infty} \frac{kM+K}{k\omega+m} = \frac{M}{\omega}.$$

Now remark again that nodes of linearly independent solutions separate each other, and we have the general theorem:

THEOREM XI. *If* (16) *is oscillatory and* $\rho_1$ *and* $\rho_2$ *are real, then the ratio of the number of nodes of any real solution, not identically zero, on an interval* $a \leqslant x < b$, *to the length of the interval approaches a limit, other than zero, as the length of the interval becomes infinite.*

Now assume $\rho_1$ and $\rho_2$ imaginary.

Let $y(i)$ be a solution of (16) such that

$$y(i+\omega) = \rho_1 y(i);$$

then
$$y(i) = u(i)+\sqrt{(-1)}v(i),$$

where $u$ and $v$ are real and linearly independent. They will be considered as continuous functions of $x$, according to the definition previously given: that is, as those functions defined by the broken-line graphs.

Choose some point $a$, and let

$$y(a) = Re^{\phi\sqrt{(-1)}},$$

where $0 \leqslant \phi < 2\pi$. When $x$ varies continuously from $a$ to $a+\omega$, $y$ changes continuously from $y(a)$ to $\rho_1 y(a)$. It never passes through the origin, as $u$ and $v$ are linearly independent and hence do not vanish together. Let $\psi$ denote the actual increment of the angle of $y$, determined by continuity during this process of variation,

$$\rho_1 = |\rho_1|e^{\psi\sqrt{(-1)}}.$$

Now $\rho_1$ is imaginary, and hence $\psi \neq 0$.

When $x$ goes from $a+(n-1)\omega$ to $a+n\omega$ the increment added to the angle of $y(x)$ by continuity is again $\psi$. Suppose it were $\psi_n = \psi+2M_n\pi$, $M_n$ an integer not equal to zero. $y\{i+(n-1)\omega\}$ is a solution of equation (16), as is readily seen by substitution. When $x$ goes from $a$ to $a+\omega$ the angle of $y\{x+(n-1)\omega\}$ determined by continuity is increased by $\psi_n$. That is, the point $y\{x+(n-1)\omega\}$ has made at least one more (or less) complete revolution around the origin than has $y(x)$. In other words, the solution $u\{x+(n-1)\omega\}$ has had at least two more (or less) nodes than $u(x)$. This is impossible, as the nodes of two solutions either coincide or separate each other. Hence we conclude

$$\psi_n = \psi.$$

Let the angle of $y(x)$ be designated by $\phi(x)$. $\phi(x)$ *either always increases or always decreases with* $x$; for let $\phi'(x)$, $u'(x)$, $v'(x)$ represent the forward derivatives of the respective functions with respect to $x$:

$$\phi'(x) = \frac{v(x)u'(x)-u(x)v'(x)}{\{u(x)\}^2+\{v(x)\}^2}.$$

This is always defined and preserves the same sign.†

We have just observed that $\psi_n$, determined by continuity as $x$ goes from $a+(n-1)\omega$ to $a+n\omega$ is independent of $n$. It follows that the number of complete revolutions of $y(x)$ around the origin as $x$ goes from $a+(n-K)\omega$ to $a+n\omega$ is also independent of $n$.

Let $$x-a = n\omega+\theta\omega, \quad 0 \leqslant \theta < 1.$$

Let $M$ be the number of complete revolutions of $y(x)$ as $x$ goes from $a$ to $a+n\omega+\theta\omega$, $N$ the number of complete revolutions of $y$ as $x$ goes from $a$ to $a+n\omega$, and let $\psi = \alpha 2\pi$; then

$$N \leqslant M < N+\alpha+2 \quad \text{and} \quad n\alpha-1 < N \leqslant n\alpha.$$

Hence $$\lim_{x\to\infty}\frac{M}{x-a} = \lim_{x\to\infty}\frac{N}{x-a} = \frac{\alpha}{\omega}.$$

Let the number of nodes of $u$ on the interval $x-a$ be $P$; then

$$2M \leqslant P \leqslant 2M+2.$$

† As we know, when $u(i)$ and $v(i)$ are linearly independent solutions of (16), the determinant $v(i)\Delta u(i)-u(i)\Delta v(i)$ never vanishes and preserves the same sign for all values of $i$. If one assumes that $v(x)u'(x)-u(x)v'(x)$ vanishes or changes sign, one can readily deduce a contradiction to this theorem.

Hence
$$\lim_{x\to\infty}\frac{P}{x-a}=\frac{2\alpha}{\omega}.$$
Since the nodes of linearly independent solutions separate each other, if $u(x)$ is any real solution of (16), not identically zero, and $\bar{P}$ the number of its nodes on $x-a$, then
$$\lim_{x\to\infty}\frac{\bar{P}}{x-a}=\frac{2\alpha}{\omega}.$$

We can summarize the foregoing results in the following two theorems:

THEOREM XII. *If $\rho_1$ and $\rho_2$ are imaginary* (16) *is oscillatory.*

THEOREM XIII. *If equation* (16) *is oscillatory the ratio of the number of nodes of any real solution, not identically zero on an interval, to the length of the interval, approaches a limit other than zero as the length of the interval becomes infinite. This limit is the same for all solutions.*

## EXERCISES

1. Prove the differential equation
$$\frac{d^2y}{dx^2}+(\mu^2+\tfrac{1}{4})\frac{y}{x^2}=0,\quad \mu^2>0,$$
to be oscillatory over the infinite interval.

Prove
$$\frac{d^2y}{dx^2}+\frac{1}{4}\frac{y}{x^2}=0$$
non-oscillatory over the infinite interval.

2. Given
$$\Delta^2y(i)+p(i)y(i+1)=0,\tag{18}$$
where $p(i)$ is defined, $a\leqslant i<\infty$, and $\lim_{i\to\infty}p(i)=0$. Prove, if $\lim_{i\to\infty}i^2p(i)>\frac{1}{4}$, then (18) is oscillatory; if $\lim_{i\to\infty}i^2p(i)<\frac{1}{4}$ or if $\lim_{i\to\infty}i^2p(i)=\frac{1}{4}$ with the approach from below, then (18) is non-oscillatory.

## SUGGESTED REFERENCES

BÔCHER, M.: 'Non-Oscillatory Linear Differential Equations of the Second Order', *Bull. Amer. Math. Soc.* **7**, 333.

FORT, T.: 'Oscillatory and Non-Oscillatory Linear Difference Equations of the Second Order', *Quart. Jour. of Pure and Ap. Math.* No. 179, 240.

KNESER, A.: 'Untersuchungen über die reellen Nullstellen der integrale linear Differentialgleichungen', *Math. Ann.* **42**, 400.

LEVI-CIVITA, M. T.: 'Sur les équations linéaires à coefficients périodiques', *Ann. Sci. de L'École N. Sup.* **28**, 325.

# XVII

# THE LINEAR DIFFERENCE EQUATION IN A CONTINUOUS REAL INDEPENDENT VARIABLE

## 1. The gamma function

THE gamma function plays such a fundamental role in the theory of differences that some discussion of it should be given at this point. However, the most interesting properties of gamma of $x$ lie in the theory of analytic functions of a complex variable. Consequently, inasmuch as this subject is omitted from the present volume the discussion that is given is perforce brief.

We give the following definition

$$\Gamma(x) = \int_0^\infty e^{-u}u^{x-1}\,du, \quad x > 0. \tag{1}$$

The integral is readily proved convergent. It can be thrown into other forms by means of various transformations. One frequently employed is the following: Let

$$u = \log\frac{1}{v}$$

and we get
$$\Gamma(x) = \int_0^1 \left(\log\frac{1}{v}\right)^{x-1} dv. \tag{2}$$

Integration by parts shows that

$$\Gamma(x+1) = x\Gamma(x). \tag{3}$$

If $x = 1$ we see from (1) that $\Gamma(1) = 1$. Consequently from (3) if $n$ is a positive integer,

$$\Gamma(n) = (n-1)!. \tag{4}$$

From the theory of integrals $\Gamma(x)$ is continuous in $x$, $x > 0$.

Formula (3) is a homogeneous linear difference equation of the first order and $\Gamma(x)$ is a continuous function, $x > 0$, which is a solution of difference equation (3). Note that the range of the independent variable $x$ is the continuum and not a discrete set of values as was the case in Chapter VII. Now equation (3) can be made a basis of definition for $\Gamma(x)$ for negative values of $x$

other than $-1, -2, -3, \ldots$, the values being successively determined by (3) from the values in the interval

$$0 < x < 1. \tag{5}$$

$\Gamma(x)$ is not defined when $x = 0, -1, -2, -3, \ldots$.

Formula (3) also makes it unnecessary to tabulate $\Gamma(x)$ for values of $x$ outside interval (5), inasmuch as their computation from the values when $x$ is in (5) is a simple matter.

We can express $\Gamma(x)$ as an infinite product. To do this we note first that

$$e^{-u} = \lim_{n\to\infty}\left(1-\frac{u}{n}\right)^n$$

and

$$\Gamma(x) = \lim_{n\to\infty}\int_0^n e^{-u}u^{x-1}\,du, \quad x > 0.$$

Then

$$\Gamma(x) = \lim_{n\to\infty}\left[\int_0^n \left(1-\frac{u}{n}\right)^n u^{x-1}\,du + \int_0^n \left\{e^{-u}-\left(1-\frac{u}{n}\right)^n\right\}u^{x-1}\,du\right].$$

Consider the second integral

$$\int_0^n \left[e^{-u}-\left(1-\frac{u}{n}\right)^n\right]u^{x-1}\,du = \int_0^n \left[1-e^u\left(1-\frac{u}{n}\right)^n\right]e^{-u}u^{x-1}\,du.$$

We wish to show that this approaches zero when $n \to \infty$. We remark the following relations: When $0 \leqslant u \leqslant n$,

$$1-\frac{u}{n} \leqslant e^{-u/n}, \tag{6}$$

$$1+\frac{u}{n} \leqslant e^{u/n}, \tag{7}$$

$$1-\frac{u^2}{n} \leqslant \left(1-\frac{u^2}{n^2}\right)^n. \tag{8}$$

The relations are all equalities if $u = 0$, and for other values of $u$ and $n$ the derivative of the left-hand member with reference to $u$ is less than that of the right-hand member. From (6), (7), and (8) we conclude that

$$e^u\left(1-\frac{u}{n}\right)^n \leqslant 1$$

and $$e^u\left(1-\frac{u}{n}\right)^n \geqslant \left(1+\frac{u}{n}\right)^n\left(1-\frac{u}{n}\right)^n \geqslant 1-\frac{u^2}{n}.$$

From these we conclude that

$$0 \leqslant 1-e^u\left(1-\frac{u}{n}\right)^n \leqslant \frac{u^2}{n}.$$

Consequently,

$$\int_0^n \left[1-e^u\left(1-\frac{u}{n}\right)^n\right]e^{-u}u^{x-1}\,du \leqslant \frac{1}{n}\int_0^n u^2e^{-u}u^{x-1}\,du \leqslant \frac{1}{n}M,$$

which approaches zero when $n \to \infty$. Hence

$$\Gamma(x) = \lim_{n\to\infty}\int_0^n \left(1-\frac{u}{n}\right)^n u^{x-1}\,du.$$

Setting $u = nv$ this gives

$$\Gamma(x) = \lim_{n\to\infty} n^x \int_0^1 (1-v)^n v^{x-1}\,dv. \tag{9}$$

Integrating by parts we get

$$\begin{aligned}\int_0^1 (1-v)^n v^{x-1}\,dv &= \frac{n}{x+n}\int_0^1 (1-v)^{n-1}v^{x-1}\,dv \\ &= \frac{n}{x+n}\,\frac{n-1}{x+n-1}\int_0^1 (1-v)^{n-2}v^{x-1}\,dv \\ &\quad \cdots\cdots\cdots\cdots \\ &= \frac{n}{x+n}\,\frac{n-1}{x+n-1}\cdots\frac{1}{x+1}\,\frac{1}{x}.\end{aligned}$$

From (9) we can thus write

$$\Gamma(x) = \lim_{n\to\infty}\frac{1}{x}\,\frac{1.2...(n-1)}{(x+1)(x+2)...(x+n-1)}\,n^x. \tag{10}$$

Now

$$n^x = \left(\frac{2}{1}\cdot\frac{3}{2}\cdot\frac{4}{3}\cdots\frac{n}{n-1}\right)^x = \left(1+\frac{1}{1}\right)^x\left(1+\frac{1}{2}\right)^x\cdots\left(1+\frac{1}{n-1}\right)^x.$$

Also

$$(x+1)(x+2)...(x+n-1) = [(n-1)!]\left(1+\frac{x}{1}\right)\left(1+\frac{x}{2}\right)\cdots\left(1+\frac{x}{n-1}\right).$$

Putting these in (10) gives

$$\Gamma(x) = \frac{1}{x}\prod_{n=1}^{\infty}\frac{(1+1/n)^x}{1+x/n}$$
$$= \frac{1}{x}\prod_{n=1}^{\infty}\frac{e^{x\log(1+1/n)}}{1+x/n}$$
$$= \frac{1}{x}\prod_{n=1}^{\infty}\frac{e^{x\log(1+1/n)}e^{-x/n}}{(1+x/n)e^{-x/n}}$$
$$= \frac{(1/x)\prod\limits_{n=1}^{\infty}e^{x[\log(1+1/n)-1/n]}}{\prod\limits_{n=1}^{\infty}(1+x/n)e^{-x/n}}.$$

It is not difficult to show that both infinite products converge for all values of $x$ other than $0, -1, -2,\ldots$. Hence

$$\Gamma(x) = \frac{1}{x}\frac{\prod\limits_{n=1}^{\infty}\exp[x\{\log(n+1)-\log n-1/n\}]}{\prod\limits_{n=1}^{\infty}(1+x/n)e^{-x/n}}$$
$$= \frac{1}{x}\frac{\exp\Big\{x\Big[\sum\limits_{n=1}^{\infty}\{\log(n+1)-\log n\}-\sum\limits_{n=1}^{\infty}1/n\Big]\Big\}}{\prod\limits_{n=1}^{\infty}(1+x/n)e^{-x/n}}.$$

Thus
$$\Gamma(x) = \frac{1}{x}\frac{e^{-\gamma x}}{\prod\limits_{n=1}^{\infty}(1+x/n)e^{-x/n}}, \tag{11}$$

where
$$\gamma = \sum_{n=1}^{\infty}\left[\frac{1}{n}-\log\left(1+\frac{1}{n}\right)\right], \tag{12}$$

is Euler's constant.

## 2. Further formulae

If in (11) we replace $x$ by $-x$ we find by multiplication that

$$\frac{1}{\Gamma(x)}\frac{1}{\Gamma(-x)} = -x^2\prod_{n=1}^{\infty}\left(1-\frac{x^2}{n^2}\right);$$

but we know that

$$\sin\pi x = \pi x\prod_{n=1}^{\infty}\left(1-\frac{x^2}{n^2}\right).$$

Hence
$$\frac{1}{-x\Gamma(x)\Gamma(-x)} = \frac{\sin\pi x}{\pi}.$$

But
$$-x\Gamma(-x) = \Gamma(1-x).$$

Hence
$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin\pi x}. \tag{13}$$

This is an interesting formula and permits the computation of $\Gamma(x)$ when $\frac{1}{2} < x \leqslant 1$ if $\Gamma(x)$ is known when $0 < x \leqslant \frac{1}{2}$.

We have developed an asymptotic form for $\log(x!)$ in Chapter IV. By only small modifications of the work there we have, when $x > 1$,

$$\log\Gamma(x+1) = -x+(x+\tfrac{1}{2})\log x + \log\sqrt{(2\pi)}+ + \sum_{r=1}^{n-1}\frac{B_{2r}}{2r(2r-1)}\frac{1}{x^{2r-1}}+R_n(x), \tag{14}$$

$$R_n(x) = \theta\frac{B_{2n}}{(2n-1)2n}x^{-2n+1},\quad 0<\theta<1. \tag{15}$$

Details are left to the reader.

From (11) we have

$$\log\Gamma(x) = -\gamma x-\log x+\sum_{n=1}^{\infty}\left[\frac{x}{n}-\log\left(1+\frac{x}{n}\right)\right].$$

Differentiating formally, we get

$$\frac{d}{dx}\log\Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)} = -\gamma+\sum_{n=1}^{\infty}\left(\frac{1}{n}-\frac{1}{x+n-1}\right).$$

This series converges uniformly in $x$ if $x$ is bounded away from $0, -1, -2,\ldots$, and hence represents the derivative.

## 3. The psi function

We give the following definition:

$$\psi(x,h) = \log x-h\sum_{j=0}^{\infty}\left[\frac{1}{x+jh}-\underset{h}{\Delta}\log(x+jh)\right]. \tag{16}$$

We readily show, by applying the law of the mean to

$$\underset{h}{\Delta}\log(x+jh),$$

that the series occurring in this definition converges uniformly

in $x$, when $x > \delta > 0$. It consequently defines a continuous function when $x > 0$. The most interesting thing about this function is the following relation

$$\underset{h}{\Delta}\psi(x, h) = \frac{1}{x}. \tag{17}$$

In other words, $\psi(x, h)$ is a continuous function of $x$ given by a comparatively simple formula and satisfies the difference equation

$$y(x+h)-y(x) = \frac{h}{x}.$$

We are, of course, familiar with other similar functions. We have called attention to the gamma function. Numerous functions satisfying the equation

$$\underset{h}{\Delta}y(x) = f(x)$$

for various particular specifications of $f(x)$ are determined by formulae of Chapter I.

The function $\psi(x, h)$ is of use in mathematical analysis and is of considerable historical interest. A few additional things about it will be developed. From (16)

$$\psi(x, h) = \log x - \sum_{j=0}^{\infty}\left[\frac{1}{(x/h)+j} - \log\left(1+\frac{1}{(x/h)+j}\right)\right]. \tag{18}$$

Now if $\gamma$ is Euler's constant,

$$\gamma = \sum_{j=0}^{\infty}\left[\frac{1}{1+j} - \log\left(1+\frac{1}{1+j}\right)\right].$$

Hence from (18)

$$\psi(x, h)+\gamma-\log h = -\sum_{j=0}^{\infty}\left(\frac{h}{x+jh} - \frac{1}{1+j}\right)+S,$$

where

$$S = \log\frac{x}{h} + \sum_{j=0}^{\infty}\left[\log\left(1+\frac{1}{(x/h)+j}\right) - \log\left(1+\frac{1}{1+j}\right)\right]$$

$$= \log\frac{x}{h} + \lim_{m\to\infty}\log\frac{(x/h)+m+1}{m+2} - \log\frac{x}{h} = 0.$$

Consequently,

$$\psi(x, h) = \log h - \gamma - \sum_{j=0}^{\infty} \left(\frac{h}{x+hj} - \frac{1}{1+j}\right).$$

This is a very convenient form for $\psi(x, h)$. We have immediately the formula

$$\frac{\partial}{\partial x}\psi(x, h) = \sum_{j=0}^{\infty} \frac{h}{(x+hj)^2},$$

a series which converges uniformly in $x$ for $x > \delta > 0$, thus assuring us that term-by-term differentiation is allowable, $x > 0$. In general

$$\frac{\partial^n}{\partial x^n}\psi(x, h) = \sum_{j=0}^{\infty} \frac{(-1)^{n-1}(n!)h}{(x+hj)^{n+1}}.$$

The functions $\psi(x, h)$, $(\partial/\partial x)\psi(x, h)$, $(\partial^2/\partial x^2)\psi(x, h)$, etc., are frequently known as digamma of $x$, trigamma of $x$, tetragamma of $x$, etc. This formula permits ready development of $\psi(x, 1)$ into a Fourier series for any interval $x_0 < x < x_0+1$. We have

$$\psi(x, 1) = \log x_0 + 2\sum_{n=1}^{\infty} [(\mathrm{ci}\, 2\pi n x_0)\cos 2n\pi x + (\mathrm{si}\, 2\pi n x_0)\sin 2n\pi x],$$

where

$$\mathrm{ci}\, x = -\int_x^{\infty} \frac{\cos t}{t}\, dt, \qquad \mathrm{si}\, x = -\int_x^{\infty} \frac{\sin t}{t}\, dt.$$

Formula (18) also permits of the obtaining of an asymptotic form for $\psi(x, h)$. The similarity to the work for $\sum_{i=1}^{n} (1/i)$ given in Chapter IV is so great, however, that details are left to the reader.

## 4. The sum of a function

The problem of obtaining analytic solutions of difference equations is one that has interested many mathematicians. Usually it is a problem of difficulty and in addition is one of primary interest in the complex domain. The gamma function and psi function are examples of such functions. The last is a solution of a difference equation of the type

$$\frac{1}{h}\Delta y(x) = \phi(x). \tag{19}$$

Other solutions of equations of this type are obtained in Chapter II. We shall proceed to a further brief consideration of (19), following Nörlund.

If the series $\sum_{j=0}^{\infty} \phi(x+jh)$ converges it is immediate that

$$F(x, h) \equiv \sum_{j=0}^{\infty} \phi(x+jh)$$

is a solution of (19). Frequently, however, this series diverges. With this in mind we set up

$$F(x, h) = \lim_{\mu \to 0} \left[ \int_c^{\infty} \phi(t) e^{-\mu\lambda(t)}\, dt - h \sum_{j=0}^{\infty} \phi(x+jh) e^{-\mu\lambda(x+jh)} \right],$$

which defines a function of $x$ and $h$, assuming that a function $\lambda(x)$ can be found so that both the integral and series converge and the expressed limit exists uniformly in $x$ over the interval considered.

THEOREM. *If the function* $F(x, h)$ *as defined above exists it is a solution of* (19).

The function in the bracket is clearly a solution of

$$\frac{1}{h} \Delta y(x) = \phi(x) e^{-\mu\lambda(x)}.$$

The theorem follows immediately from the fact that the difference of the limits is the limit of the difference.

The solution which we have obtained depends upon an arbitrary constant $c$. We shall call it the principal solution of (19). Following Nörlund we write

$$F(x, h) = \mathop{S}_{c}^{x} \phi(t) \Delta t.$$

There are certain analogies to an integral. The general solution of (19) is given by adding to $F(x, h)$ an arbitrary function of period $h$.

We give two elementary examples; in the first the introduction of the convergence factors $e^{-\lambda(t)\mu}$ is not necessary. Consider

$$\frac{1}{h} \Delta y(x) = e^{-x}.$$

Here

$$F(x, h) = \int_c^\infty e^{-t}\,dt - h\sum_{j=0}^\infty e^{-x-jh} = e^{-c} - \frac{he^{-x}}{1-e^{-h}}.$$

For a second example consider

$$\frac{1}{h}\Delta y(x) = a,$$

where $a$ is a constant. Take $\lambda(x) = x$:

$$F(x, h)$$

$$= \lim_{\mu\to 0}\left[\int_c^\infty ae^{-\mu t}\,dt - h\sum_{j=0}^\infty ae^{-\mu(x+jh)}\right]$$

$$= \lim_{\mu\to 0}\frac{ae^{-\mu c}\left[\left(\mu h - \frac{1}{2!}\mu^2h^2+...\right)-\{\mu h - \mu^2h^2(x-c)+...\}+...\right]}{\mu\left(\mu h - \frac{\mu^2h^2}{2!}+...\right)}$$

$$= a(x-c-\tfrac{1}{2}h).$$

For $h = 1$, $c = 0$, $a = 1$ we have $x-\frac{1}{2} \equiv B_1(x)$; thus the Bernoulli polynomial is a principal solution.

The subject of the 'sum of a function' has been extensively studied by Nörlund and Milne-Thomson. Only the smallest introduction is given here. Again the topic properly belongs in the study of the difference equation in the complex domain and for that reason no more extensive study is made here.

## 5. The linear equation of the $n$th order

A linear difference equation of the $n$th order with continuous independent variable is an equation of the form†

$$p_0(x)y(x+n)+p_1(x)y(x+n-1)+...+p_n(x)y(x) = r(x), \quad (20)$$

where $p_0(x), p_1(x),..., p_n(x), r(x)$ are defined when

$$a \leqslant x \leqslant b; \quad (21)$$

and $p_0(x)p_n(x) \neq 0$ at any‡ point of the interval (21).

A solution of (20) is any function of the continuous variable $x$

† There is no gain in generality if the difference interval is taken as $h$ instead of 1. Simply change the variable, letting $x = ht$.

‡ Less restrictive conditions on $p_0(x)$ and $p_n(x)$ than this can be imposed. Modifications of the ensuing theory must naturally be made.

which, if substituted for $y$, satisfies the equation at all points of (21). The solution is necessarily defined over the interval

$$a \leqslant x \leqslant b+n. \tag{22}$$

As a matter of fact it is customary to call a function a solution of (20) even if there are sets of congruent isolated points at which it is not defined, provided that the equation is satisfied at all points at which the function is defined. The term particular solution is used to describe a uniquely defined solution and the term general solution to describe a formula including all solutions as special cases.

Many things about equation (20) follow immediately from the corresponding theory of equation (2) of Chapter VII.

We have the following existence theorem:

THEOREM. *Let $c = a+m$, where $m$ is zero or a positive integer and such that $c+n < b$. There exists one and only one solution of (20), defined when $a \leqslant x \leqslant b+n$, such that*

$$\begin{aligned} y(x) &= f_1(x), \quad c \leqslant x < c+1, \\ y(x) &= f_2(x), \quad c+1 \leqslant x < c+2, \\ &\ldots\ldots\ldots\ldots \\ y(x) &= f_n(x), \quad c+n-1 \leqslant x < c+n, \end{aligned}$$

*where $f_1(x), f_2(x),..., f_n(x)$ are arbitrary single-valued functions.*

The solution $y(x)$ is determined by the successive use of (1). Such a solution will in general be discontinuous and not given by a compact and usable formula.

Equation (20) is called homogeneous if $r(x) \equiv 0$. It is called non-homogeneous if $r(x) \not\equiv 0$. For reference we write the homogeneous equation, after dividing through by $p_0(x)$,

$$y(x+n)+P_0(x)y(x+n-1)+...+P_n(x)y(x) = 0. \tag{23}$$

Let $y_1(x),..., y_m(x)$ be particular solutions of (23). They are said to be *linearly dependent* over (22) if there exist single-valued functions $c_1(x),..., c_m(x)$ with period 1, not all zero at any point, such that

$$c_1(x)y_1(x)+c_2(x)y_2(x)+...+c_n(x)y_n(x) \equiv 0. \tag{24}$$

They are said to be *linearly independent* in the contrary case.

If $y_1(x),..., y_m(x)$ are solutions of (23) they are said to form a

fundamental system of solutions if every solution is expressible in the form

$$y(x) \equiv c_1(x)y_1(x)+c_2(x)y_2(x)+...+c_m(x)y_m(x),$$

where $c_1(x)$, $c_2(x)$,..., $c_m(x)$ have the period 1, and if this is true of no sub-set.

THEOREM. *Any n linearly independent solutions of* (23) *constitute a fundamental system of solutions and every fundamental system consists of exactly n linearly independent solutions.*

THEOREM. *A necessary and sufficient condition that the solutions of* (23), $y_1,..., y_n$, *be linearly independent is*

$$W(y_1,..., y_n) \equiv \begin{vmatrix} y_1(x) & y_2(x) & \dots & y_n(x) \\ y_1(x+1) & y_2(x+1) & \dots & y_n(x+1) \\ \dots & \dots & \dots & \dots \\ y_n(x+n-1) & y_2(x+n-1) & \dots & y_n(x+n-1) \end{vmatrix} \neq 0 \quad (25)$$

*at any point of* (21).

THEOREM. *The n functions* $y_1(x)$, $y_2(x)$,..., $y(x)$, *linearly independent over the interval* $a \leqslant x \leqslant b$ *in the sense of the definition just given, serve to determine uniquely a linear homogeneous difference equation of the n-th order for which they constitute a fundamental system of solutions.*

Substitute the given functions successively in equation (23) and there results a system of linear algebraic equations for the determination of the coefficients which possess a unique solution due to the non-vanishing of the determinant $W(y_1,..., y_n)$.

THEOREM. *The general solution of the non-homogeneous equation* (20) *consists of the general solution of the corresponding homogeneous equation* (23) *plus any particular solution of* (20).

## 6. The equation of the first order

We consider $$y(x+1)-A(x)y(x) = B(x) \quad (26)$$

and the corresponding reduced equation

$$y(x+1)-A(x)y(x) = 0, \quad (27)$$

where $A(x)$ and $B(x)$ are defined, $a \leqslant x \leqslant b$, and $B(x) \not\equiv 0$.

Denote by $\alpha_x$ the point of the interval $a \leqslant x < a+1$ which is such that $x-\alpha_x$ is an integer. Let $f(x)$ be an arbitrary single-valued function defined $a \leqslant x \leqslant a+1$. Then

$$y(x) \equiv f(\alpha_x)A(\alpha_x)A(\alpha_x+1)...A(x-1) \tag{28}$$

is the general solution of (27). Similarly, if

$$u(x) = A(\alpha_x)A(\alpha_x+1)...A(x-1),$$

$$y = u(x)\sum \frac{B(x)}{u(x+1)}$$

is the general solution of (26).

A typical equation of the type (8) is

$$y(x+1)-xy(x) = 0.$$

## 7. The linear equation with constant coefficients

Procedure here is as with the equation with integral argument only. Whenever explicit forms for the solution can be found, simply replace the variable $i$ by the variable $x$ throughout.

## 8. Difference equations which can be reduced to linear equations with constant coefficients

The equation

$$\begin{gathered} y(x+n)+A_1\phi(x)y(x+n-1)+A_2\phi(x)\phi(x-1)y(x+n-2)+ \\ +...+A_n\phi(x)\phi(x-1)...\phi(x-n+1)y(x) = r(x), \qquad (29) \\ \phi(x) \neq 0, \end{gathered}$$

can be reduced to an equation with constant coefficients by the substitution

$$y(x+j) = \phi(x+j-n)...\phi(\alpha_x)u(x+j).$$

The transformed equation is

$$u(x+n)+A_1 u(x+n-1)+...+A_n u(x) = \frac{r(x)}{\phi(x)\phi(x-1)...\phi(\alpha_x)}.$$

Here for any value of $x$ we choose an initial value $\alpha_x$ such that $x-\alpha_x$ is an integer.

The equation

$$y(x+1)y(x)+a(x)y(x+1)+b(x)y(x) = c(x)$$

can be reduced to a linear equation of the second order and

under certain conditions to a linear equation with constant coefficients. Make the substitution

$$y(x) = \frac{v(x+1)}{v(x)} - a(x),$$

and reduce. We obtain

$$v(x+2)+\{b(x)-a(x+1)\}v(x+1)-\{a(x)b(x)+c(x)\}v(x) = 0.$$

This equation may be an equation with constant coefficients or one which can be so reduced by the method just discussed.

There are a variety of other devices which may reduce a non-linear equation to a linear equation or which will reduce the equation to some familiar form. The matter is in every way similar to the analogous problem in differential equations. We give one more illustration. Others will be considered in the exercises.

Consider

$$(a+bx)\Delta^2 y(x)+(c+dx)\Delta y(x)+ey(x) = 0, \tag{30}$$

where $-e/d = n$ is a positive integer. Perform the operation $\Delta^n$ on (30). We obtain

$$[a+b(x+n)]\Delta^{n+2}y(x)+[nb+c+d(x+n)]\Delta^{n+1}y(x) = 0.$$

This is an equation of the first order in $\Delta^{n+1}y(x)$ and consequently can be completely solved first for $\Delta^{n+1}y(x)$ and then for $y(x)$. This last operation will introduce constants of summation (periodic functions) which must be determined by substitution in the given equation.

## 9. Multipliers

Consider a difference equation

$$L(x) = l_0(x)\Delta^n y(x)+l_1(x)\Delta^{n-1}y(x)+\ldots+l_n(x)y(x) = 0, \tag{31}$$

where $l_0(x)$, $l_1(x)$,..., $l_n(x)$ are defined for all values of $x$. In case they are originally defined over a finite interval only, such as

$$a \leqslant x \leqslant b,$$

we define them for other values of $x$ in any arbitrary manner. It is desired to find a function $v(x)$ such that

$$v(x)L(x) = \Delta[m_0(x)\Delta^{n-1}y(x)+m_1(x)\Delta^{n-2}y(x)+\ldots+m_{n-1}(x)y(x)], \tag{32}$$

where $m_0(x), m_1(x),..., m_{n-1}(x)$ are defined for all values of $x$. To do this we expand the right-hand member in (32) and equate coefficients. We have

$$\begin{aligned} v(x)l_0(x) &= m_0(x+1), \\ v(x)l_1(x) &= \Delta m_0(x)+m_1(x+1), \\ &\cdots\cdots\cdots \\ v(x)l_{n-1}(x) &= \Delta m_{n-2}(x)+m_{n-1}(x+1), \\ v(x)l_n(x) &= \Delta m_{n-1}(x). \end{aligned} \tag{33}$$

We eliminate $m_0, m_1,..., m_{n-1}$ as follows. Take the difference of both sides of the first of equations (33). Increase the argument in the second equation by 1 and subtract. Take the difference of the resulting expression. Increase the argument in the third equation by 1 and subtract. Proceeding in this manner we obtain the following equation

$$\Delta^n[l_0(x)v(x)]-\Delta^{n-1}[l_1(x+1)v(x+1)]+\Delta^{n-2}[l_2(x+2)v(x+2)]- \\ -...+(-1)^n l_n(x+n)v(x+n) = 0. \tag{34}$$

Equation (34) has been obtained as a necessary condition for (32). It is, however, also sufficient since a function $v(x)$ which satisfies (34) permits of the solution of (33) for $m_0(x),..., m_{n-1}(x)$.

Equation (34) is called the *adjoint* equation to (31). The function $v(x)$ is called a *multiplier*.

## 10. Linear $q$-difference equations

Equations of the following type are commonly known as linear $q$-difference equations:

$$a_0(x)y(x)+a_1(x)y(qx)+a_2(x)y(q^2x)+...+a_n(x)y(q^{n-1}x) = b(x).$$

Here $q$ is a constant different from 1.

Many studies of this equation have been made. Particularly is this true in the field of analytic functions. The student is referred to a paper by C. R. Adams,† where a summary of this work is given with a very complete bibliography.

† *Bull. Amer. Math. Soc.* **37**, 361–400.

## EXERCISES

Solve the following equations:

1. $y = x\Delta y + (\Delta y)^2$;
2. $(\Delta y)^2 - 5\Delta y + 6 = 0$;
3. $y(x+1) - 4.3^{2x}y(x) = 5.3^{x^2}$;
4. $y(x+1) - 4y(x) = \cos nx$;
5. $y(x)y(x+1) + 6y(x) + 7 = 0$;
6. $y(x+1) - e^{2x-1}y(x) = e^{x^2}$;
7. $y(x+1) - ay(x) = (2x+1)a^x$;
8. $y(x+1) = m\{y(x)\}^n$;
9. $\Delta^2 y(x) = [y(x+1)]^2 - [y(x)]^2$;
10. $y(x+2) - 9y(x+1) + 20y(x) = e^x - \sin x$;
11. $y(x+2) - 2(x-1)y(x+1) + (x-1)(x-2)y(x) = x$;
12. $(x+3)^2 y(x+2) - \dfrac{2(x+2)^2}{x+1} y(x+1) + \dfrac{(x+1)^2(x+2)}{x} y(x) = 0$;
13. $y(x+1) - v(x) = 2m(x+1)$,
    $v(x+1) - y(x) = -2m(x+1)$;
14. $y(x+2) + 2v(x+1) - 8y(x) = a^x$,
    $v(x+2) - y(x+1) - 2v(x) = a^{-x}$.

15. Form the homogeneous linear difference equation of the second order for which $y_1 = 3x+2$ and $y_2 = 2x-1$ constitute a fundamental system of solutions.

16. Show that
$$\int\limits_0^\infty x^m e^{-px^q}\,dx = \frac{\Gamma\{(m+1)/q\}}{qp^{(m+1)/q}}, \quad m > -1, \quad p > 0,\ q > 0.$$

17. Show that
$$\int\limits_0^1 x^n(\log x)^m\,dx = \frac{(-1)^m\Gamma(m+1)}{(m+1)^{m+1}}, \quad \begin{matrix} m > -1, \\ n > -1. \end{matrix}$$

18. Show that
$$\int\limits_0^1 \frac{(-\log x)^p}{1-x}\,dx = \Gamma(p+1)\sum_{k=1}^\infty \frac{1}{k^{p+1}}, \quad p > 0.$$

19. Show that $\quad \Gamma(2x) = \dfrac{1}{\sqrt{\pi}} 2^{2x-1}\Gamma(x)\Gamma(x+\tfrac{1}{2})$.

20. By definition $\quad B(x, y) = \dfrac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$.

Show that
$$\int\limits_0^\infty \frac{x^m}{(1+px^q)^k}\,dx = \frac{1}{qp^{(m+1)/q}} B\Big(\frac{m+1}{q}, k - \frac{m+1}{q}\Big),$$
$$p > 0, \quad q > 0, \quad m > -1, \quad k > \frac{m+1}{q}.$$

**21.** Show that

$$\int_0^p x^m(p^q-x^q)^n\,dx = \frac{p^{qn+m+1}}{q}\,B\!\left(n+1,\frac{m+1}{q}\right),$$

$$p>0,\quad q>0,\quad m>-1,\quad n>-1.$$

**22.** Show that

$$\psi(x,1) = \log x - \sum_{j=0}^{\infty}\left[\frac{1}{x+j}-\log\!\left(1+\frac{1}{x+j}\right)\right].$$

**23.** Show that

$$\psi(x,1) = \log x - \frac{1}{2x} - \sum_{j=0}^{\infty}\left[\frac{2(x+j)+1}{2(x+j)(x+j+1)}-\log\frac{x+j+1}{x+j}\right].$$

**24.** Show that $\quad \psi(x,1) = -\gamma+\int_0^1 \frac{1-t^{x-1}}{1-t}\,dt.$

**25.** Show that

$$\psi(x,1) = \log x - \frac{1}{2x} + 2\int_0^x \frac{t\,dt}{(x^2+t^2)(1-e^{2\pi t})}.$$

# BOOKS FOR PARALLEL READING


BLEICH, FR., and MELAN, E.: *Die gewöhnlichen und partiellen Differenzengleichungen der Baustatik.* Berlin, Springer, 1927.

BÖHMER, P. E.: *Differenzengleichungen und bestimmte Integrale.* Leipzig, 1939.

BOOLE, G.: *A Treatise on the Calculus of Finite Differences.* Reprinted, New York, Stechert, 1926.

JORDAN, C.: *Calculus of Finite Differences.* Budapest, Rottig & Romwalter, 1939.

MARKOFF, A. A.: *Differenzenrechnung.* German translation from Russian by Th. Friesendorff and H. Prumm, Leipzig, 1896.

MILNE-THOMSON, L. M.: *The Calculus of Finite Differences.* London, Macmillan, 1933.

NÖRLUND, N. E.: *Vorlesungen über Differenzenrechnung.* Berlin, Springer, 1924.

SELIWANOFF, D.: *Lehrbuch der Differenzenrechnung.* Leipzig, 1904.

STEFFENSEN, J. F.: *Interpolation.* Baltimore, Williams & Wilkins, 1927.

WALLENBERG, G., and GULDBERG, A.: *Theorie der linearen Differenzengleichungen.* Berlin, Teubner, 1911.

# INDEX

PRINTED IN
GREAT BRITAIN
AT THE
UNIVERSITY PRESS
OXFORD
BY
CHARLES BATEY
PRINTER
TO THE
UNIVERSITY